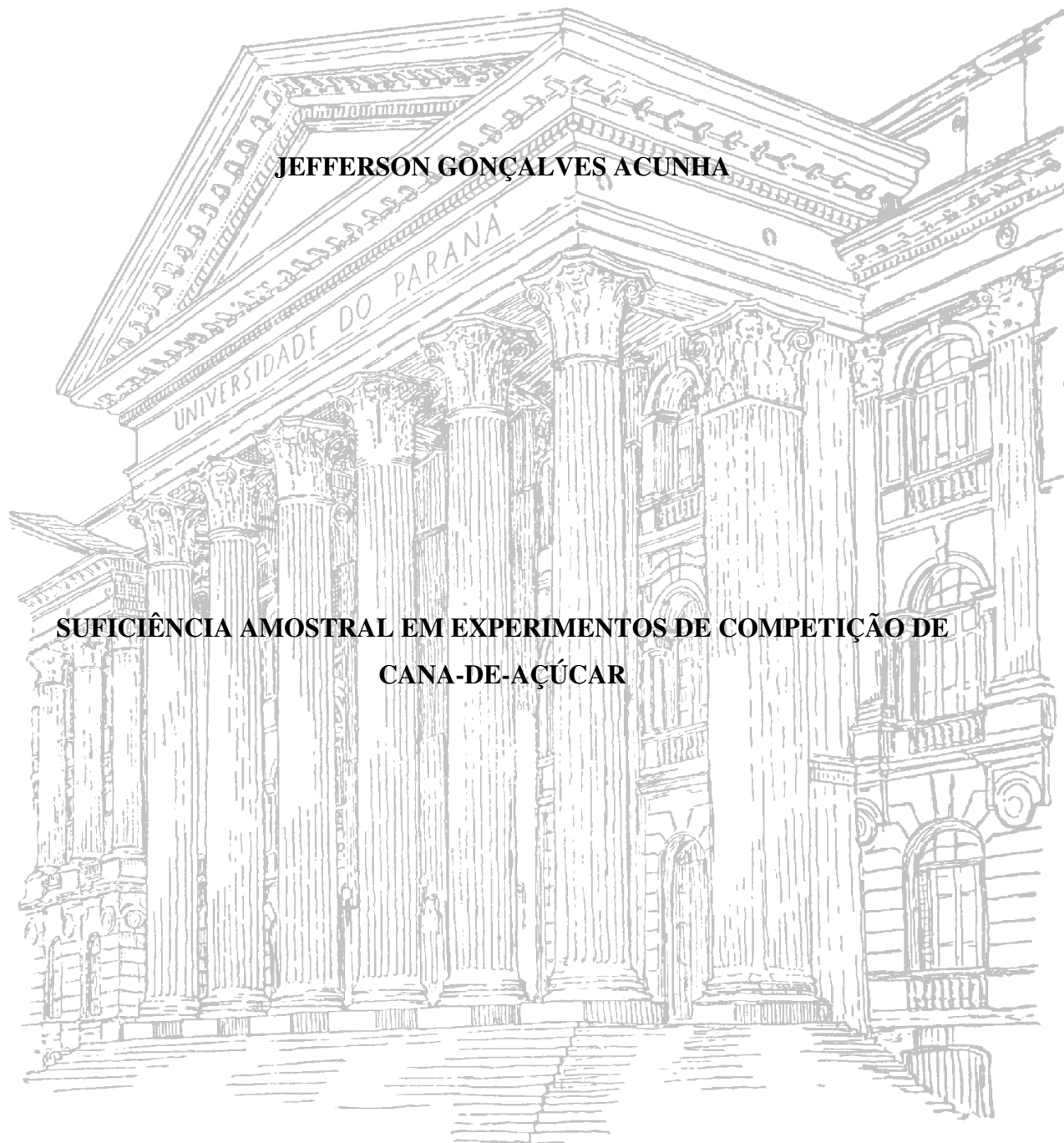


UNIVERSIDADE FEDERAL DO PARANÁ

JEFFERSON GONÇALVES ACUNHA

**SUFICIÊNCIA AMOSTRAL EM EXPERIMENTOS DE COMPETIÇÃO DE
CANA-DE-AÇÚCAR**



CURITIBA

2013

JEFFERSON GONÇALVES ACUNHA

SUFICIÊNCIA AMOSTRAL EM EXPERIMENTOS DE COMPETIÇÃO DE
CANA-DE-AÇÚCAR

Dissertação apresentada como requisito parcial à obtenção do título de Mestre em Ciências, Área de Concentração em Produção Vegetal, no Curso de Pós-Graduação em Agronomia, Setor de Ciências Agrárias, da Universidade Federal do Paraná.

Orientador: Prof. Dr. Ricardo Augusto de Oliveira
Co-orientador: Prof. Dr. Walmes Marques Zeviani

CURITIBA
2013

A189 Acunha, Jefferson Gonçalves
 Suficiência amostral em experimentos de competição de cana-de-açúcar.
 / Jefferson Gonçalves Acunha. – Curitiba, 2013.
 77 f. : il.

Orientador: Ricardo Augusto de Oliveira
Co-orientador: Walmes Marques Zeviani
Dissertação (Mestrado em Agronomia) – Universidade Federal do Paraná. Setor de Ciências Agrárias. Programa de Pós-Graduação em Agronomia – Produção Vegetal

1. Cana-de-açúcar – Melhoramento genético.
2. Cana-de-açúcar – Amostragem (Estatística). I. Oliveira, Ricardo Augusto de. II. Zeviani, Walmes Marques. III. Universidade Federal do Paraná. Setor de Ciências Agrárias. Programa de Pós-Graduação em Agronomia – Produção Vegetal. IV. Título

CDU 633.61:631.528.6



UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS AGRÁRIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM
AGRONOMIA - PRODUÇÃO VEGETAL



PARECER

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Agronomia - Produção Vegetal, reuniram-se para realizar a arguição da Dissertação de MESTRADO, apresentada pelo candidato **JEFFERSON GONÇALVES ACUNHA**, sob o título "**SUFICIÊNCIA AMOSTRAL EM EXPERIMENTOS DE COMPETIÇÃO DE CANA-DE-AÇÚCAR**", para obtenção do grau de Mestre em Ciências do Programa de Pós-Graduação em Agronomia - Produção Vegetal do Setor de Ciências Agrárias da Universidade Federal do Paraná.

Após haver analisado o referido trabalho e argüido o candidato são de parecer pela "**APROVAÇÃO**" da Dissertação.

Curitiba, 21 de Maio de 2013.

Professora Dra. Louise Larissa May De Mio
Coordenadora do Programa

Dr. Edilson Batista de Oliveira
Primeiro Examinador

Professor Dr. Walmes Marques Zeviani
Segundo Examinador

Professor Dr. João Carlos Bessalho Filho
Terceiro Examinador

Professor Dr. Ricardo Augusto de Oliveira
Presidente da Banca e Orientador

AGRADECIMENTOS

Ao meu orientador e ao meu co-orientador, Prof. Dr. Ricardo Augusto de Oliveira e Prof. Dr. João Carlos Bessalho Filho, por terem me oportunizado a realização deste trabalho, bem como pelo seu constante apoio e atenção.

Ao professor Walmes Marques Zeviani, do Departamento de Estatística da Universidade Federal do Paraná, pelos inestimáveis ensinamentos relativos ao programa “R”.

Ao professor Dr. Luiz Alexandre Peternelli, do Departamento de Estatística da Universidade Federal de Viçosa, por ter-se disponibilizado a colaborar na execução deste trabalho, gentilmente.

Aos colegas Mário Álvaro A. Veríssimo e Michael Jonathan F. Alves, doutorandos, bem como ao acadêmico Daniel Sousa, pelo seu imprescindível auxílio nas atividades de campo.

A todo o pessoal do Programa de Melhoramento Genético da Cana-de-açúcar (RIDESA), pelos amistosos e frutíferos momentos de convivência, bem como pelo suporte oferecido na coleta dos dados experimentais.

A todos os professores e colegas, tanto da Universidade Federal de Santa Maria, quanto da Universidade Federal do Paraná, que têm contribuído para a minha formação integral.

À secretaria e ao colegiado do Programa de Pós-graduação em Agronomia – Produção Vegetal, por terem compreendido as dificuldades pelas quais passei ao longo deste mestrado.

Evidentemente, a Deus, o criador de tudo e de todos, na pessoa de Jesus Cristo, o seu filho unigênito, meu “patrão” eterno, amigo e defensor, pelas bênçãos concedidas a mim, das quais jamais serei digno.

RESUMO

A cana-de-açúcar é uma das principais culturas agrícolas do Brasil e do mundo, sendo fonte ímpar de alimento e bioenergia. Nas etapas do melhoramento desta cultura há grande demanda por recursos humanos e materiais. Para a avaliação dos experimentos da fase experimental, que corresponde aos ensaios de competição e indicação de cultivares, tem-se recomendado a colheita total de parcelas, algo que, devido aos seus custos, tem-se mostrado inexecutável. Alternativamente, pode-se estimar a produtividade com base na tomada de amostras que representem as parcelas. O objetivo deste trabalho foi estudar a suficiência amostral em cana-de-açúcar para a estimação da massa, do teor de sólidos solúveis de parcelas experimentais. Também procurou-se definir os tamanhos de amostra suficientes à avaliação experimental dos teores de sacarose da cultura. Reamostragens com reposição subsidiaram o uso dos modelos da curva potencial, linear e quadrático segmentados com resposta platô e do modelo do coeficiente de variação relativo. Também foram realizados estudos simulados de análise de variância. Os resultados indicaram que amostras de 20 colmos coletados aleatoriamente na parcela experimental permitem estimar adequadamente a variável massa do colmo. Com este valor, serão obtidas avaliações confiáveis da produtividade em colmos (TCH). Para o BRIX, leituras em 10 colmos de 10 touceiras aleatoriamente coletadas são suficientes à sua estimativa. Para as avaliações experimentais de sacarose (POL), podem ser usados feixes que contenham ao menos 20 colmos.

Palavras-chave: Cana-de-açúcar, tamanho da amostra, suficiência amostral, melhoramento da cana-de-açúcar, *Saccharum* spp.

SAMPLE SUFFICIENCY IN COMPETITION EXPERIMENTS WITH SUGARCANE

ABSTRACT

Sugarcane is one of the most important crops, both in Brazil as in the world, being an unmatched source of food and bioenergy. Sugarcane breeding demands human and material resources in large-scale. During the evaluation of the experimental phase, which corresponds to competition experiments and indication of cultivars, is recommended total harvest of plots, which, owing to their costs, can become impracticable. Optionally, one can estimate the productivity based on taking samples representing plots. The aim of this study was to research the sample sufficiency in sugarcane for estimating the plot mass and the brix contents of experimental plots. One tried to define also the sufficient sample size for the experimental evaluation of the sucrose contents. One used the maximum curvature of the power model, the linear and quadratic segmented models with response plateau and the relative CV model. Studies based on simulated analysis of variance were done also. The application of these models and the ANOVA was supported by bootstrap resamplings. The results indicated that a sample made by 20 stalks harvested in plot can be used for estimating properly the variable mass of the stalk. This value may be used for estimating the sugarcane yield per hectare. For brix variable, 10 stalks of 10 different plants are sufficient for estimating this variable. For POL, bundles containing at least 20 stalks can be sampled for estimating this variable. Samples with such dimensions can be used to estimate appropriately brix yield and sucrose yield per hectare.

Keywords: Sugarcane, sample size, sample sufficiency, sugarcane breeding, *Saccharum* spp.

LISTA DE FIGURAS

FIGURA 1: Exemplo de curva potencial, destacando-se o seu ponto de máxima curvatura.	27
FIGURA 2: Esboço ilustrativo da obtenção do ponto de máxima curvatura da função potencial pelo método de Barros e Tavares (1995).	32
FIGURA 3: Esboço do modelo linear segmentado com platô.	33
FIGURA 4: Esboço do modelo quadrático com resposta platô.....	34
FIGURA 5: Gráfico explicativo do modelo do CV relativo, tal qual utilizado por Carneiro <i>et al.</i> (2012).....	36
FIGURA 6: Gráficos dos ajustes dos modelos CP, LSP, QRP e CVR, para uma parcela (bloco um) da cultivar padrão. As linhas tracejadas remetem às abscissas dos pontos correspondentes à suficiência amostral obtida através de cada modelo, para a variável MC.....	45
FIGURA 7: Gráficos dos ajustes dos modelos CP, LSP, QRP e CVR, para uma parcela (bloco um) da cultivar padrão. As linhas tracejadas remetem às abscissas dos pontos correspondentes à suficiência amostral obtida através de cada modelo, para a variável BRIX.....	46
FIGURA 8: Gráfico do ajuste dos modelo CP aos dados de CV exp. da variável POL em função do tamanho da amostra. A linha tracejada remete à abscissa do ponto correspondente à suficiência amostral obtida pelo MCM II.....	51
FIGURA 9: Para cada tamanho de amostra, proporções dos p-valores significativos e dos <i>rankings</i> cujos valores da correlação de Spearman foram maiores do que 0,9.	52

LISTA DE TABELAS

TABELA 1: Valores de R^2 e de AIC (Critério de Informação de Akaike) para os modelos CP, LSP, QRP e CVR ajustados às variáveis MC e BRIX.....	47
TABELA 2: Dimensões da amostra ideal para as variáveis MC e BRIX, segundo os métodos MCM I, MCM II, LSP, QRP e CVR. Todos os resultados foram arredondados para o maior valor inteiro subsequente.....	48
TABELA 3: CV experimentais – variável POL – obtidos pelas ANOVA executadas para cada tamanho de amostra e comparações das suas estimativas aos valores relatados por Couto <i>et al.</i> (2013).	50

LISTA DE ANEXOS

ANEXO 1 – QUADROS DE ANOVA RESUMIDOS	65
ANEXO 2 – ALGUNS ALGORITMOS DE PROGRAMAÇÃO UTILIZADOS NO AMBIENTE COMPUTACIONAL “R”	66
ANEXO 3 – GRÁFICOS RESULTANTES DOS <i>BOOTSTRAPS</i> E DOS AJUSTES DOS MODELOS A TODAS AS PARCELAS	71

SUMÁRIO

RESUMO	6
ABSTRACT	7
1. INTRODUÇÃO	12
2. REVISÃO DE LITERATURA	15
2.1 A CULTURA DA CANA-DE-AÇÚCAR	15
2.2 COMPETIÇÃO DE CLONES DE CANA-DE-AÇÚCAR	18
2.3 SUFICIÊNCIA AMOSTRAL E AMOSTRAGEM EM CANA-DE-AÇÚCAR.....	18
2.4 METODOLOGIAS PARA ESTIMAR-SE A SUFICIÊNCIA AMOSTRAL.....	21
2.5 REAMOSTRAGEM <i>BOOTSTRAP</i>	23
2.6 DESCRIÇÃO TEÓRICA DE ALGUMAS METODOLOGIAS DE DIMENSIONAMENTO AMOSTRAL.....	24
2.6.1 Análise de Variância.....	24
2.6.2 Intensidade de Amostragem	25
2.6.3 Máxima Curvatura da Curva Potencial	27
2.6.4 Modelos Segmentados com Resposta Platô	32
2.6.5 Modelo da Curva de Michaelis-Menten Reparametrizada ou Modelo do CV Relativo (“ <i>Relative CV Model</i> ”)	35
3. CAPÍTULO – SUFICIÊNCIA AMOSTRAL EM EXPERIMENTOS COM CANA-DE- AÇÚCAR	37
3.1 RESUMO.....	37
3.2 ABSTRACT.....	38
3.3 INTRODUÇÃO	39
3.4 MATERIAIS E MÉTODOS.....	40
3.5 RESULTADOS E DISCUSSÃO.....	45
3.6 CONCLUSÕES	53
3.7 CONSIDERAÇÕES FINAIS	53
3.8 REFERÊNCIAS.....	54
4. REFERÊNCIAS GERAIS.....	58
ANEXO 1 – QUADROS DE ANOVA RESUMIDOS.....	65
ANEXO 2 – ALGUNS ALGORITMOS DE PROGRAMAÇÃO UTILIZADOS NO AMBIENTE COMPUTACIONAL “R”.....	66
ANEXO 3 – GRÁFICOS RESULTANTES DOS <i>BOOTSTRAPS</i> E DOS AJUSTES DOS MODELOS A TODAS AS PARCELAS.....	71

1. INTRODUÇÃO

Fornecedora de sacarose, etanol e biomassa energética, a cana-de-açúcar (*Saccharum* spp.), no Brasil atual, é cultivada numa área de mais de oito milhões e trezentos mil hectares, tendo atingido, na safra 2012, uma produção de 595,13 milhões de toneladas (CONAB, 2012). Pelas suas características geográficas e edafo-climáticas, o Brasil tem-se mantido como maior produtor mundial de cana, cujo cultivo vem seguindo em relativa expansão.

Para que a agroindústria sucroalcooleira seja cada vez mais competitiva, continue atendendo às crescentes demandas energéticas e alimentares, há que se disponibilizar frequentemente novos genótipos aos produtores, buscando-se, desta forma, maiores níveis de produtividade e melhor gerenciamento dos riscos inerentes a toda atividade agropecuária. Durante o processo de melhoramento genético da cana-de-açúcar, muitos experimentos são conduzidos, a fim de que os melhoristas possam, de fato, selecionar adequadamente genótipos superiores. Este processo, de maneira resumida, inicia-se com a reprodução sexuada, a partir de cruzamentos controlados (nos quais pelo menos um parental deverá ser conhecido). Após esta fase, tendo-se obtido as cariopses (frutos secos indeiscentes que contêm as sementes botânicas), são produzidos os *seedlings*, os quais serão submetidos aos primeiros processos de seleção, o que irá constituir a fase T1 de seleção. As seleções prosseguirão pelas fases ditas T2 e T3, em cana planta e soca (e ressoça, às vezes), sendo que, a partir desta última fase, os clones promissores serão multiplicados e distribuídos a um número maior de estações experimentais, a fim de que se possa passar à “Fase Experimental” de seleção clonal, cujo propósito principal seria o de se obter, dentre os melhores clones, o(s) superior(es), propícios a serem lançados como variedades (cultivares) comerciais (BARBOSA e SILVEIRA, 2000; e BRESSIANI, 2001).

À fase final de seleção, podem ser classificados em torno de 30 clones de ciclo precoce e 30 de ciclo tardio, os quais serão comparados mediante experimentos conduzidos em vários locais e sob o delineamento de “blocos casualizados” (doravante, DBC), com no mínimo três repetições, gerando-se, assim, grande demanda por áreas experimentais e enorme volume de trabalho, o qual, normalmente, encontra-se sob a responsabilidade de equipes com restrições orçamentárias e de mão de obra. Assim, a otimização e o bom uso dos recursos disponíveis às avaliações destes ensaios, constitui-

se num singular desafio, cujas respostas passam, indubitavelmente, pela busca por metodologias experimentais economicamente viáveis e cientificamente corretas, a fim de que os programas de melhoramento genético possam ter continuidade, disponibilizando, de fato, os melhores materiais às usinas e fornecedores.

A operação agrícola fundamental à avaliação da produtividade dos experimentos agrônômicos é a colheita das parcelas experimentais, a qual pode ser realizada de modo pleno, ou seja, colhendo-se a totalidade das parcelas (das suas áreas úteis, quando da utilização de bordaduras), ou de modo amostral, sendo ambas portadoras de vantagens e desvantagens. Dentre as vantagens da colheita total, deve-se destacar a possibilidade de se obter o valor real da produção de cada parcela, que será usado para a estimação da produtividade por área. Obviamente, esta metodologia de colheita concorre para a majoração dos custos operacionais, além de ser, no geral, altamente demandante em termos de mão de obra, maquinário e insumos. Por sua vez, optando-se pela colheita parcial (por amostragem), há que se levar em conta o fato de que a extrapolação da produtividade da parcela para áreas maiores (um hectare, consuetudinariamente), será baseada em duas estimativas, a saber: a primeira, da produção total pela amostra; e a segunda, da estimativa de produtividade da parcela para outra área. Este problema de sobre-estimativa inerente à amostragem, pode ser enfrentado através da validação de metodologias que indiquem tanto a suficiência amostral (o tamanho ideal de amostra) para a avaliação de uma determinada variável, quanto um procedimento de amostragem exequível para a obtenção da amostra. Tendo-se abordado adequada e cientificamente estas duas questões, pode-se desfrutar de todas as vantagens advindas da colheita amostral de parcelas, a saber: economia de tempo necessário à execução das operações (colheita e pós-colheita de material, em alguns casos); redução da necessidade de mão de obra operacional; economia de insumos; e redução de custos oriundos da aquisição e manutenção de máquinas e implementos.

Devido ao grande número de experimentos a campo, à sobreposição temporal das diversas fases experimentais, à demanda por recursos humanos e materiais e aos custos de mecanização das operações agrícolas em cana-de-açúcar, é necessário que se busque validar a colheita amostral desta cultura, proporcionando aos programas de melhoramento genético da cana-de-açúcar subsídios à melhor gestão dos recursos disponíveis, tanto do ponto de vista econômico quanto do ponto de vista agrônômico.

Diante do exposto, este trabalho buscará validar a suficiência amostral para os experimentos da “Fase Experimental” (doravante, FE) em cana-de-açúcar. Para tal,

serão utilizados procedimentos estatísticos já consagrados de experimentação agrícola aplicada ao fitomelhoramento, a fim de se definir o tamanho ideal de amostra que represente tanto a massa de colmos (MC) das parcelas experimentais com cana-de-açúcar quanto o seu teor de sólidos solúveis (BRIX). Também será estudada a suficiência amostral para a correta avaliação do teor de sacarose (POL) e da produtividade em colmos (TCH) nos experimentos com cana-de-açúcar. No âmbito deste objetivo, serão avaliados os impactos dos tamanhos de amostra indicados pelo estudo da suficiência amostral, para a variável MC, nas estimativas do índice TCH. Por fim, a partir do estudo das variáveis supracitadas e do índice TCH, discutir-se-á a sua aplicabilidade direta às estimativas dos índices de produtividade derivados TBH (toneladas de BRIX por hectare) e TPH (toneladas de POL por hectare).

2. REVISÃO DE LITERATURA

2.1 A CULTURA DA CANA-DE-AÇÚCAR

A cana-de-açúcar (*Saccharum* spp.) é uma gramínea tropical perene, também adaptada aos subtrópicos, pertencente ao gênero *Saccharum* da tribo *Andropogoneae*, subfamília *Panicoideae*, da família botânica *Poaceae* (CHEAVEGATTI-GIANOTTO *et al.*, 2011). Fisiologicamente, trata-se de uma planta cuja rota fotossintética é a C₄, sendo, portanto, altamente eficiente na fixação de carbono, desde que cresça sob condições favoráveis às espécies C₄, ou seja, sob ótima disponibilidade de radiação solar e temperaturas de medianas a altas (KERBAUY, 2008). Parte significativa dos seus fotossintatos podem ser acumulados como sacarose de reserva no seus colmos, tornando-a uma espécie energética por excelência.

Modernamente, o termo “cana-de-açúcar”, refere-se a inúmeras cultivares (variedades), melhoradas ou não, pertencentes a um complexo botânico chamado “complexo *Saccharum*”, no qual estão incluídos, além daquelas variedades de *Saccharum* spp., outros gêneros correlatos, como *Mischantus* e *Erianthus* (JAMES, 2004).

Acredita-se que o centro de origem das duas espécies que mais têm contribuído, mediante hibridação, para a constituição do genoma das variedades atuais sejam o subcontinente indiano, para *S. spontaneum* (mas há controvérsias quanto a isto), e a região de fronteira entre o sudeste asiático e a Oceania (CHEAVEGATTI-GIANOTTO *et al.*, 2011). De fato, há indícios de que os primeiros atos de domesticação da cana-de-açúcar tenham ocorrido na Nova Guiné e nos arquipélagos polinésios e melanésios próximos, conforme James (2004), havendo grande probabilidade de que *S. officinarum* tenha se originado de seleções efetuadas em *S. robustum* (JAMES, 2004). Por outro lado, na Índia, cruzamentos naturais com *S. barberi* teriam originado as mais antigas variedades daquele subcontinente, as quais devem ter sido matéria-prima para a primeira agroindústria açucareira mundial, ao redor do ano 3000 a.C., de acordo com Cesnik e Miocque (2004). Da Índia, o cultivo da cana foi levado à Pérsia (Irã) e deste país ao mundo mediterrâneo, pelos árabes, que já dominavam as técnicas de extração do açúcar (CESNIK e MIOCQUE, 2004). Após esta etapa, coube aos navegadores ibéricos

a expansão da cultura para outros locais (insulares e continentais), então pertencentes ao seu domínio.

No Brasil, a sua introdução oficial (pois há relatos extra-oficiais) se deu por Martim Afonso de Sousa, donatário da Capitania de São Vicente, o qual portava mudas oriundas da Ilha da Madeira, território ultramarino português (CESNIK e MIOCQUE, 2004). Atualmente, cultiva-se a cana em todo o território nacional, merecendo destaque, contudo, as regiões canavieiras do centro-sul (geograficamente composta por alguns estados das regiões sudeste, centro-oeste e sul) e nordestina. Os três maiores estados produtores são os seguintes (CONAB, 2012): 1º) São Paulo, que concentra 51,87% da área cultivada com cana no Brasil; 2º) Goiás, 8,52% da área total; e 3º) Minas Gerais, com 8,47% do total. Pelas suas dimensões territoriais, diversificação agropecuária e pelo fato de boa parte da sua área estar sob um regime climático de invernos frios, com grande ocorrência de geadas (clima do tipo “Cfb”), o Paraná merece ser destacado como grande produtor de cana, açúcar e energia, ocupando a quarta colocação no *ranking* nacional, em termos de área plantada e de produção total (CONAB, 2012).

Na base da cadeia agrícola sucroenergética, podem ser alocados os programas de Melhoramento Genético da cultura, responsáveis não apenas pelo desenvolvimento de novas variedades, voltadas a inúmeras classes de produtores (grandes produtores, como no caso das usinas e fornecedores, médios e pequenos produtores de açúcar, melado, rapadura, cana para garapa, cachaça, etc., e pequenos pecuaristas), mas também pela conservação, introdução e intercâmbio de germoplasma, algo fundamental para evitar-se que se trabalhe sobre uma base genética muito estreita e carente de variabilidade, mesmo no caso de uma planta de genoma poliplóide como a cana-de-açúcar.

O trabalho de melhoramento genético costuma ser dividido em várias etapas e, no caso da cultura da cana, costuma-se dividi-lo, resumidamente, nas seguintes fases (BARBOSA e SILVEIRA, 2000):

- Cruzamentos ou hibridação: Trata-se da etapa inicial, realizadas em estações localizadas em regiões propícias ao florescimento da cultura, onde, em determinadas épocas, são efetuados os cruzamentos solicitados pelos melhoristas (programados). O final desta fase corresponde à colheita das cariopses (que contêm as sementes botânicas, sementes verdadeiras, da cana-de-açúcar), nas quais, através da reprodução sexuada, novos genótipos são obtidos (segregação gênica). Os frutos de cada cruzamento efetuado

constituirão uma família de indivíduos geneticamente aparentados (irmãos germanos, meio-irmãos, etc.).

- Fase T1: Está relacionada ao cultivo e à avaliação dos *seedlings* oriundos da germinação daquelas cariopses. A seleção de plantas tem início nesta fase, podendo ser “massal”, realizada sobre os indivíduos, conforme a descrição de Cesnik e Miocque (2004), ou por meio da seleção de famílias (em Bressiani (2000) e Oliveira (2007) podem ser obtidas informações detalhadas acerca desta fase e das suas metodologias de seleção). Trata-se da fase com os mais baixos índices de seleção, por se tratar de progênies oriundas de populações segregantes, ficando tal índice ao redor dos 5% (segundo Cesnik e Miocque (2004), na prática, os índices costumam ficar bem abaixo deste valor).
- Fase T2: Nesta fase, de acordo com Cesnik e Miocque (2004), os clones começam a ser comparados com clones padrões. São avaliadas em cana planta e soca uma série de características morfofisiológicas e fitossanitárias, além dos componentes do rendimento. O índice de seleção, para esta fase, ficaria em torno de 10%. O delineamento experimental utilizado nesta fase é o de blocos aumentados, também chamado de blocos de Federer (MATSUOKA *et al.*, 1999).
- Fase T3: As avaliações desta fase, cujo delineamento é o mesmo da fase anterior, referem-se a aspectos produtivos (quilogramas de BRIX e POL por parcela) e fitossanitários (tolerância e susceptibilidade a determinadas doenças, como ferrugens, carvão, mosaico e podridão vermelha), sendo considerada uma importante fase complementar à fase T2, em termos de seleção, conforme Matsuoka *et al.* (1999).
- FM (fase de multiplicação): Trata-se de uma fase intermediária que serve para a multiplicação vegetativa de material (obtenção de mudas) a ser plantados nos ensaios da fase seguinte, a fase experimental de seleção, conhecida pela sigla FE.
- FE (fase experimental): Esta é a fase final de um programa de melhoramento, na qual os clones promissores são avaliados em ensaios de competição de clones promissores, tendo como objetivo prever-se o lançamento de novas variedades (CESNIK e MIOCQUE, 2004). Além disto,

são conduzidos experimentos visando à construção da curva de maturação específica de cada variedade, a fim de se indicar o seu período útil de industrialização (PUI).

2.2 COMPETIÇÃO DE CLONES DE CANA-DE-AÇÚCAR

Os experimentos de competição de clones de cana-de-açúcar correspondem à última fase dentre aquelas que compõem um programa de melhoramento genético da cultura. Os clones costumam ser avaliados quanto às suas variáveis produtivas (TCH, TPH, etc.), tecnológicas (POL, teor de fibras e de BRIX, etc.) e quanto à sua responsividade ambiental (BARBOSA e SILVEIRA, 2000; CESNIK e MIOCQUE, 2004; MATSUOKA *et al.*, 1999).

Nesta fase, os experimentos são normalmente instalados em vários locais, sendo avaliados em distintos ambientes de produção, conforme o DBC com três ou quatro blocos (BARBOSA e SILVEIRA, 2000). A condução se estende por no mínimo três ciclos produtivos, abrangendo os estádios de cana planta, primeiro rebrote ou soca, e segundo rebrote ou ressoca, enquanto que as avaliações costumam ocorrer em pelo menos três épocas de corte (MATSUOKA *et al.*, 1999). Paralelamente, são conduzidos ensaios de curva de maturação, pelos quais se busca definir o período útil de industrialização da cana (PUI) mediante a descrição do comportamento dos níveis de sacarose ao longo do tempo (CESNIK e MIOCQUE, 2004).

Os resultados das avaliações experimentais são submetidos a análises de variância e testes de comparação de médias, considerando-se os genótipos como tratamentos de efeito fixo (CRUZ *et al.*, 2004).

2.3 SUFICIÊNCIA AMOSTRAL E AMOSTRAGEM EM CANA-DE-AÇÚCAR

Para que se possa detectar a superioridade de alguma cultivar com relação a outras, de forma científica, há que se lançar mão de ferramentas próprias da

Experimentação Agrícola, subárea das ciências estatísticas e agronômicas (STORCK *et al.*, 2000).

Dentre os vários procedimentos experimentais necessários à correta condução experimental, encontra-se a obtenção de informações, ou seja, a coleta dos dados de interesse obtidos a partir de um determinado experimento (STORCK *et al.*, 2000). Contudo, devido a vários motivos (dimensões da área e das parcelas experimentais, limitações de mão de obra e/ou de recursos, limitações de tempo para a execução das operações de amostragem, impossibilidade de se mensurar a totalidade da população em estudo, etc.), nem sempre será possível coletar-se os dados referentes à totalidade de uma determinada população em estudo, sendo, portanto, necessária a tomada de amostras da população em questão (FONSECA e MARTINS, 1995). Todos os procedimentos utilizados na tomada e construção das amostras constituem a amostragem propriamente dita (SILVA, 2007). Ora, tais amostras devem ser suficientemente representativas, a fim de que as estimativas dos parâmetros obtidos através delas encontrem-se dentro de níveis de confiança satisfatórios, evitando-se conclusões equivocadas, distorcidas e enviesadas, afinal todas as inferências que se façam acerca de uma determinada população em estudo serão baseadas naquelas mesmas amostras (BALESTRASSI e PAIVA, 2007).

No processo de definição de um tamanho ideal de amostra, deve-se atentar para os seguintes fatos: a) desde que os estimadores sejam consistentes, com o aumento do número de unidades básicas amostradas, aumenta-se também a precisão das estimativas dos parâmetros que se queira conhecer (SILVA *et al.*, 1993); b) contudo, há um limite, em termos de tamanho de amostra, para este aumento significativo de precisão, a partir do qual haveria pouca vantagem, do ponto de vista econômico, em se tomar amostras de maiores dimensões (SILVA *et al.*, 1993); c) a relação entre a precisão das estimativas calculadas e a dimensão de uma amostra é diretamente proporcional, havendo entre ambas e o erro permitido pelo pesquisador uma relação inversamente proporcional (CARGNELUTTI FILHO *et al.*, 2009); e d) pode-se aumentar o tamanho da amostra, tendo-se a certeza de se aumentar a precisão experimental, enquanto o erro experimental seguir sendo significativamente maior do que o erro amostral, ou seja, enquanto a variação entre unidades básicas amostrais for maior do que a variação interparcelar existente (STORK *et al.*, 2006).

Com relação às várias etapas do processo de melhoramento da cultura da cana-de-açúcar, há poucos estudos abordando a temática da suficiência amostral, dentre os

quais, a ampla maioria tem-se concentrado nas etapas iniciais do melhoramento, especialmente na fase inicial T1, no âmbito das metodologias de seleção de famílias. Tais estudos têm tido como objetivo a determinação do número satisfatório de indivíduos a serem amostrados para estimação do teor de sólidos solúveis (BARBOSA *et al.*, 1999), de parâmetros genéticos e de produtividade (LEITE, 2007). Para além desta fase, Matsuoka *et al.* (1999) mencionam a validação, baseada nos trabalhos de Kang¹ *et al.* (1983, citados por MATSUOKA *et al.*, 1999), da pesagem de uma amostra de 10 colmos para a estimação da massa das parcelas monolineares da fase T2.

Nas avaliações da fase de competição de clones (FE), devido à carência de pesquisas que validem procedimentos biométricos aplicados à suficiência amostral, tem-se recomendado a colheita total das parcelas experimentais (BARBOSA e SILVEIRA, 2000) para a estimação do TCH, sugerindo-se como alternativa à colheita total a exclusão das linhas externas, consideradas como bordaduras (CESNIK e MIOCQUE, 2004). Em ambos os casos (deixando-se ou não bordaduras relativas às parcelas individuais), percebe-se quão onerosa e demandante, em termos de mão de obra (operários rurais), equipamentos (dinamômetros de campo de altíssima capacidade, máquinas carregadoras, etc.) e tempo operacional, pode se tornar a avaliação de extensos campos plantados com clones em FE.

Diante deste quadro, os programas de pesquisa agrônômica e de melhoramento genético da cultura da cana-de-açúcar, têm procurado aderir ao uso de procedimentos de amostragem para a avaliação da produtividade em experimentos com cana-de-açúcar (ARIZONO *et al.*, 1998). Neste sentido, Gheller *et al.* (1999) e Landell *et al.* (1999) publicaram valiosos métodos alternativos para a estimação da produtividade da cultura em substituição à colheita total de parcelas. No primeiro trabalho, os autores procuraram estabelecer uma recomendação de amostragem para a avaliação indireta da produtividade da cultura, enfocando mais os procedimentos de amostragem do que o dimensionamento amostral. Por sua vez, o segundo estudo enfatizava a avaliação de alguns métodos de amostragem que fossem aplicáveis a áreas sob colheita mecanizada.

¹ KANG, M. S.; MILLER, J. D.; TAI, P. Y. P. **Genetic and Phenotypic Path Analysis and Heritability in Sugarcane.** Crop Science, 23: 643-647.1983.

2.4 METODOLOGIAS PARA ESTIMAR-SE A SUFICIÊNCIA AMOSTRAL

Na literatura especializada, há várias metodologias descritas para a estimação do tamanho amostral ideal (suficiência amostral), não existindo, todavia, alguma que seja unanimamente aplicável à totalidade das culturas agrícolas, florestais e aos rebanhos sob experimentação, sendo necessário o estabelecimento de critérios, por parte do pesquisador, para a correta escolha de uma metodologia de estimação que melhor se adeque às características e objetivos do seu experimento e ao conjunto de dados que tenham sido obtidos através dele. Dentre tais métodos, podem ser citados os seguintes:

- i. análise de variância, através da análise da significância das estatísticas F obtidas para os tamanhos de amostra comparados (SILVA *et al.* 1998), dos valores da “eficiência das estimativas” (SILVA *et al.*, 1993), ou por meio da declaração do tamanho amostral como sendo um fator experimental com n níveis testados (CATAPATTI *et al.*, 2008);
- ii. métodos de determinação matemática da intensidade de amostragem, baseados na relação entre uma medida de variação estimada a partir dos dados populacionais e um intervalo de confiança máximo (pré-determinado) para a estatística média (COCHRAN, 1977; PÉLLICO NETO e BRENA, 1997; CARGNELUTTI FILHO *et al.*, 2008; CARGNELUTTI FILHO *et al.*, 2009; FEIJÓ *et al.*, 2006; LÚCIO *et al.*, 2003; NUNES *et al.*, 2006; SILVA *et al.*, 1997; e TOEBE *et al.*, 2011);
- iii. métodos de determinação da máxima curvatura a partir do ajuste do modelo da curva potencial aos dados de uma medida de variação em função dos tamanhos de amostra (ARAÚJO *et al.*, 2009; BARBOSA *et al.*, 1999; BARROS e TAVARES, 1995; CARNEIRO *et al.*, 2012; FEIJÓ *et al.*, 2006; e PEQUENO *et al.*, 2003);
- iv. método do ajuste de modelos segmentados com resposta platô (FERREIRA, 2011; e CARNEIRO *et al.*, 2012);

- v. método do ajuste do modelo de Michaelis-Menten reparametrizado, também denominado “*relative CV model*” – modelo do CV relativo (CARNEIRO et al., 2012);
- vi. método da amplitude de intervalos de confiança *bootstrap* “percentil” (XIE e MOSJIDIS, 1997; LEITE, 2007; CARGNELUTTI FILHO et al., 2010; e CARGNELUTTI FILHO et al., 2012).

Nas citações acima, procurou-se restringir os exemplos de trabalhos à determinação da suficiência amostral, por conta dos objetivos desta pesquisa. Outrossim, deve-se observar que as metodologias referentes aos itens iii, iv, e v, desenvolvidas para o cálculo do tamanho ideal de parcelas, blocos e de experimentos, têm sido igualmente utilizadas para cálculos de dimensionamento amostral. Originalmente, tal uso decorre da aplicação da chamada Lei Empírica de Smith (SMITH, 1938), que busca explicar a correlação existente entre parcelas adjacentes a partir da estimação de um índice de heterogeneidade do solo, a estudos relativos ao tamanho de amostras, descrevendo a variabilidade existente entre as parcelas de um experimento, de modo a subsidiar a condução de experimentos de boa precisão e confiabilidade. Esta descrição, conforme Barros e Tavares (1995), dá-se a partir da percepção da existência de uma relação empírica entre a variância (ou entre o coeficiente de variação) da parcela e as suas dimensões ($V_x = V_1/X^b$), em que V_1 seria a variância da média associada à parcela cujo tamanho corresponda a uma unidade básica, V_x representaria a variância de uma parcela de x unidades básicas, X seria o número de unidades básicas e b o índice de heterogeneidade). Apesar de haver vinculado nominalmente o seu modelo estatístico a fatores de solo, o próprio Smith (1938) admite que o seu índice poderia explicar quaisquer fatores que gerassem variação interparcelar (climáticos, de manejo, dentre outros). De forma análoga, encontra-se semelhante relação empírica entre as variâncias de uma amostra de tamanho n, em termos de unidades básicas de amostragem, e aquela referente a uma amostra unitária, o que tem estendido a aplicação desta metodologia ao cálculo da suficiência amostral (MARODIM et al., 2000; FEIJÓ et al., 2006; NUNES et al., 2006).

A aplicação das metodologias citadas nos itens iii, iv, v, e vi pressupõe o uso de técnicas matemáticas de regressão não linear, haja visto comportamento claramente não linear do decréscimo da variabilidade experimental de características sob avaliação

(como produção de colmos ou de sacarose, no caso da cana-de-açúcar, por exemplo) com relação a incrementos nos números de unidades básicas (SMITH, 1938).

As metodologias mais modernas, por dependerem, na prática, da utilização de processamento computacional intensivo (SILVA, 1995), seriam aquelas correspondentes ao item “vi”, as quais necessariamente lançam mão de procedimentos de reamostragem *bootstrap* (vide seção 3.5).

2.5 REAMOSTRAGEM *BOOTSTRAP*

Trata-se de um método de inferência estatística desenvolvido por Efron (1979), que se baseia na construção de distribuições amostrais para estatísticas de interesse, mediante procedimentos iterativos de reamostragem com reposição (FOX e WEISBERG, 2012).

As técnicas *bootstrap* consideram que, diante de poucas informações acerca da população em estudo (o seu padrão de distribuição, por exemplo), pode-se considerar a amostra como se esta fosse “a própria população”, funcionando como a melhor fonte disponível de dados para a realização de inferências estatísticas, de modo que “*the population is to the sample as the sample is to the bootstrap samples*”² (FOX e WEISBERG, 2012). Sendo assim, uma das grandes vantagens dos procedimentos de *bootstrap*, para além da sua relativa simplicidade, é a sua relativa independência com relação à distribuição dos dados.

Oyeyemi (2008) descreve, de modo sucinto, os seguintes passos para a aplicação do *bootstrap*:

- 1º Passo: De uma conjunto de dados (amostra) original composto pelos elementos $x_1, x_2, x_3, x_i, \dots, x_n$, toma-se n amostras com reposição $s_1, s_2, s_3, s_i, \dots, s_n$, de maneira que cada elemento reamostrado, após ter sido utilizado na composição de qualquer reamostra, seja repostado ao conjunto de dados, ficando disponível para uma subsequente reamostragem.

² A população está para a amostra, tal qual a amostra está para as amostras oriundas do *bootstrap* (tradução nossa).

- 2º Passo: Logo, para cada amostra obtida, calcula-se a estimativa para um parâmetro populacional θ desconhecido, obtendo-se n estimativas de θ ($\hat{\theta}_n$), de maneira que a estimativa *bootstrap* daquele parâmetro populacional seja dada por:

$$\hat{\theta}_B = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_i,$$

sendo,

$\hat{\theta}_B$: estimativa *bootstrap* do parâmetro em estudo; e

$\frac{1}{n} \sum_{i=1}^n \hat{\theta}_i$: média das n estimativas do tipo *bootstrap* calculadas.

Além da sua utilização para fazer-se inferências a partir de amostras, o método *bootstrap* tem sido utilizado para estimar-se o viés e o desvio-padrão de estimadores e de testes de hipóteses, e para a construção de intervalos de confiança para estatísticas de interesse (FOX e WEISBERG, 2012).

2.6 DESCRIÇÃO TEÓRICA DE ALGUMAS METODOLOGIAS DE DIMENSIONAMENTO AMOSTRAL

2.6.1 Análise de Variância

Segundo Montalván e Barbin (1999), trata-se de uma metodologia de análise desenvolvida por Ronald Aylmer Fisher, que se fundamenta na repartição da variância total observada num experimento entre diferentes fontes de variação (tratamentos, blocos, resíduo, por exemplo), de modo que se possa testar a hipótese da existência de diferenças estatísticas entre as médias experimentais devidas aos fatores controlados. Para tal finalidade, lança-se mão da estatística F.

Estimativas da variância residual determinadas pela ANOVA convencional foram utilizadas por Silva *et al.* (1993) para a comparação entre tamanhos de amostra e técnicas de amostragem na avaliação de caracteres agronômicos de milho. Neste trabalho, como um dos critérios de determinação da amostra ideal, estipulou-se um índice de eficiência relativa das amostras, definido como a estimativa da variância residual obtida na ANOVA da parcela total dividida pela estimativa da variância residual obtida na análise efetuada para cada tamanho amostral.

Silva *et al.* (1998) utilizaram a análise da significância da estatística F da ANOVA como um dos critérios de decisão para a recomendação do tamanho amostral e do método de amostragem ideais na estimação de características biométricas de espigas de milho.

Trabalhando na determinação da suficiência amostral para a avaliação de caracteres agronômicos de milho pipoca, Catapatti *et al.* (2008) declararam, em modelo de ANOVA aplicado a um DBC fatorial, o tamanho de amostra como sendo tratamento de efeito fixo, o que lhes possibilitou comparar diferentes tamanhos de amostra entre si, bem como avaliar possíveis interações entre estes e os genótipos testados.

No caso da cultura de cana-de-açúcar, Barbosa *et al.* (1999) compararam, mediante ANOVA, distintos tamanhos de amostra em experimentos correspondentes à fase T1 do melhoramento da cultura, para a determinação da suficiência amostral representativa da massa de parcelas experimentais.

2.6.2 Intensidade de Amostragem

Trata-se da determinação da suficiência amostral mediante a proposição de relações matemáticas, que relacionam estatísticas calculadas com semi-amplitudes de intervalos de confiança para a média (aceitos pelo pesquisador), para um determinado nível de significância pré-admitido. No contexto desta metodologia, destacam-se as proposições de Cochran (1977) e de Péllico Neto e Brena (1997).

Na sua relação matemática, Cochran (1977) decidiu-se pelo uso do CV, da seguinte forma:

$$n_0 = \frac{t_{\alpha/2}^2 CV^2}{D},$$

na qual,

n_0 : intensidade de amostragem (tamanho de amostra recomendado);

$t_{\alpha/2}$: valor tabelado da estatística t de Student;

CV: Coeficiente de variação, expresso em percentagem; e

D: semi-amplitude, expressa percentualmente, do intervalo de confiança para a média da variável em questão.

O valor obtido desta maneira serve para o caso de populações infinitas, havendo um ajuste proposto pelo próprio Cochran (1977), caso se esteja trabalhando com populações finitas, a saber:

$$n_f = \frac{n_0}{1 + \frac{n_0}{N}},$$

onde,

n_f : tamanho da amostra recomendado; e

N: tamanho da população (finita).

Utilizando-se desta expressão, Cargnelutti Filho *et al.* (2009) estimaram o tamanho mínimo de amostra para a avaliação de vários caracteres morfológicos de soja. Semelhante estudo foi realizado para feijão (CARGNELUTTI FILHO *et al.*, 2008). Haesbaert *et al.* (2011), lançaram mão desta metodologia para a determinação da suficiência amostral em experimentos com feijão-de-vagem. Lúcio *et al.* (2003) utilizaram-na para a determinação do tamanho de amostra para a avaliação de algumas características de pimentão cultivado em estufas plásticas.

Através de uma modificação no modelo supradescrito, na qual se substituiu o CV pela variância, determinaram-se o tamanho de amostra ideal para diversos caracteres agrônômicos de mamona (CARGNELUTTI FILHO *et al.*, 2010), o tamanho amostral suficiente à estimativa do BRIX em talhões cultivados com melão amarelo (NUNES *et al.*, 2006) e o dimensionamento amostral ideal para a avaliação da firmeza da polpa e da cor da epiderme em frutos de maçã e pêssgo (TOEBE *et al.*, 2011).

2.6.3 Máxima Curvatura da Curva Potencial

Baseando-se no modelo empírico de Smith (1938), pode-se ajustar uma curva potencial do tipo $y = ax^{-b}$ (ou $y = a/x^b$) a um conjunto de dados obtidos experimentalmente, relacionando-se o comportamento da variância, do CV, ou de outras estatísticas calculadas, a distintos (e crescentes) tamanhos amostrais (LESSMAN e ATKINS, 1963). Ao avaliar-se o comportamento da curva gerada pelo modelo (FIGURA 1), pode-se perceber que, a partir de uma determinada dimensão amostral, haveria pouca vantagem em se aumentar o tamanho da amostra, em termos de ganhos de precisão, expressos por significativos decréscimos do CV, por exemplo (BARROS e TAVARES, 1995). Geometricamente, este valor de n ideal poderia ser representado pelo ponto de máxima curvatura da curva em questão, cuja determinação tem sido executada tanto por métodos algébricos quanto por inspeção visual (PARANAÍBA *et al.*, 2009).

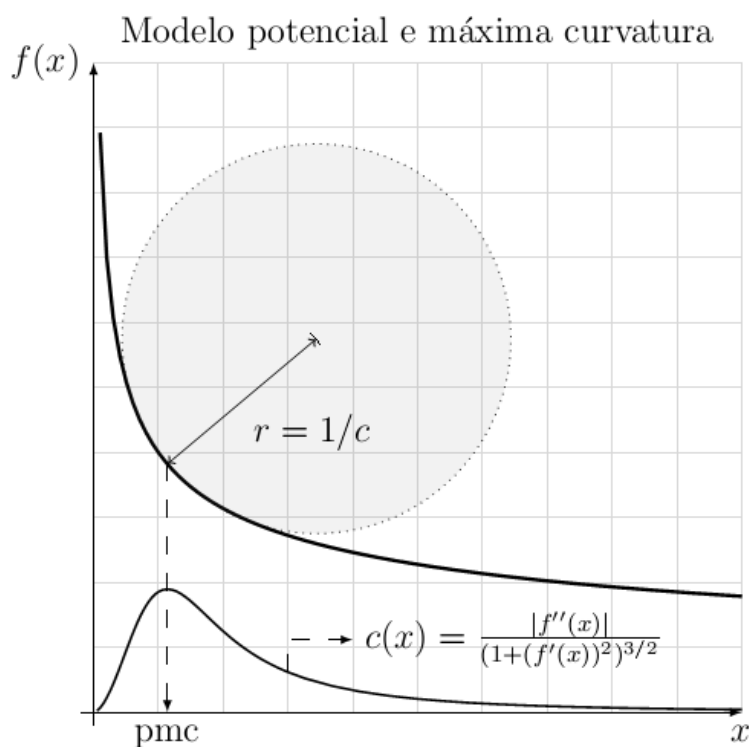


FIGURA 1: Exemplo de curva potencial, destacando-se o seu ponto de máxima curvatura.

Quanto ao método de inspeção visual, trata-se de uma tentativa expedita, prática e simples de se determinar o ponto de inflexão da curva obtida, cuja execução se dá a partir da mera observação de um gráfico que represente a curva supracitada. Por depender de inúmeros fatores subjetivos relacionados ao avaliador, esta técnica contém em si mesma, inexoravelmente, uma multiplicidade de fontes de erros e incertezas, carecendo da objetividade necessária a tais estudos (PARANAÍBA *et al.*, 2009). Além disto, Smith (1938) alertou que o posicionamento do ponto de curvatura seria alterado com qualquer mudança na escala gráfica representativa da curva, o que levaria a interpretações dúbias e discrepantes. De fato, a utilização da inspeção visual para a determinação do ponto de máxima curvatura pressupõe que o gráfico seja plotado num plano cartesiano regular, no qual a relação entre valores iguais de abscissa e ordenada seja igual a 1.

Em busca de resultados mais seguros e precisos, vários pesquisadores desenvolveram técnicas alternativas à determinação do ponto crítico de máxima curvatura da curva potencial. Estes métodos caracterizam-se por serem geométrico-algébricos, objetivos e universalmente aplicáveis, portanto. Dentre tais técnicas, destaquem-se aqui as seguintes (cronologicamente ordenadas): i) método da máxima curvatura modificado I, proposto por Meier e Lessman (1971); ii) método da máxima curvatura modificado II, proposto por Chaves³ (1985, citado por OLIVEIRA *et al.*, 2011) ; e iii) método da máxima curvatura modificado III, desenvolvido por Barros e Tavares (1995).

Método da Máxima Curvatura Modificado I

Trata-se de um método sugerido por Meier e Lessman (1971), a partir das suas críticas às proposições metodológicas de Lessman e Atkins (1963), no tocante ao cálculo da máxima curvatura. De acordo com Araújo *et al.* (2009), nesta técnica, leva-se em conta o fato de que a curvatura depende do comprimento do arco de circunferência considerado e não apenas de incrementos no tamanho amostral, o que havia servido de base ao trabalho de Lessman e Atkins (1963). Desta forma, tomando-se incrementos fixos ao longo da curva ao invés de fazê-lo ao longo do eixo das abscissas, obtém-se a

³ CHAVES, L. J. **Tamanho da Parcela para Seleção de Progenies de Milho (*Zea mays* L.)**. 1985. 148 f. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, 1985.

seguinte expressão, para o cálculo da curvatura, de modo genérico (ARAÚJO *et al.*, 2009):

$$k = \frac{y''}{\left(1 + \frac{y'}{2}\right)^{\frac{3}{2}}},$$

na qual,

k: curvatura de um ponto qualquer de $f(x)$;

y' : derivada primeira de $f(x)$; e

y'' : derivada segunda de $f(x)$.

Considerando-se $f(x) = a / x^b$, derivando-se k (com relação a x), igualando-se esta primeira derivada a zero e isolando-se a variável x, chega-se ao valor do ponto em que a curvatura seria máxima, X_{MCI} , a saber:

$$X_{MCI} = \left(\frac{a^2 b^2 (2b + 1)}{b + 2} \right)^{\left(\frac{1}{2b+2} \right)},$$

onde,

X_{MCI} : ponto de máxima curvatura da função potencial; e

a e b: parâmetros da curva.

Para a determinação do tamanho ideal de parcelas, este método tem sido usado por vários pesquisadores, os quais têm trabalhado em diversos locais e com diferentes culturas. Com relação ao dimensionamento amostral, Araújo *et al.* (2009) utilizaram este método na determinação do tamanho de amostra para a avaliação do estado de maturação de uva niágara rosada. Trabalhando com cana-de-açúcar, Leite (2007) utilizou-o para determinar a suficiência amostral para a seleção de famílias de cana-de-açúcar em experimentos da fase T1, comparando-o com outros métodos. Para determinarem um tamanho ideal de amostra que representasse o teor médio de sólidos solúveis em experimentos de avaliação de famílias de irmãos germanos de cana-de-açúcar, Barbosa *et al.* (1999) lançaram mão deste mesmo método, pelo qual obtiveram resultados compatíveis com as suas fontes bibliográficas referenciais.

Método da Máxima Curvatura Modificado II

Corresponde a uma pequena modificação aplicada ao modelo anteriormente citado, desenvolvida por Chaves (1985, citado por OLIVEIRA *et al.*, 2011), o qual sugeriu que se utilizasse o simétrico do valor de “b” da fórmula proposta por Meier e Lessman (1971), na sequência de cálculo do ponto de máxima curvatura. Mediante esta modificação, a expressão da qual se obteria o valor da abscissa do ponto de máxima curvatura “X_{MCI}” seria a seguinte:

$$X_{MCI} = \left(\frac{a^2 b^2 (2b - 1)}{b - 2} \right)^{\left(\frac{1}{2 - 2b} \right)},$$

em que,

X_{MCI}: ponto de máxima curvatura da função potencial; e

a e b: parâmetros da curva.

Método da Máxima Curvatura Modificado III

Barros e Tavares (1995) apresentaram este método como sendo mais uma alternativa à técnica de inspeção visual (FIGURA 2). Na sequência, serão apresentadas, de forma resumida e adaptada, as principais etapas para a determinação do ponto de máxima curvatura da curva potencial, tal qual foram descritas por BARROS e TAVARES (1995), a saber:

- 1º Passo: Determinam-se os valores dos parâmetros a e b da curva potencial $y = a / x^b$, que representa o comportamento da estatística *y* versus o tamanho amostral *x*. Registra-se a equação potencial ajustada (EQUAÇÃO 1), passando-se ao passo seguinte.

$$y = a / x^b \text{ (EQUAÇÃO 1);}$$

- 2º Passo: Define-se uma reta que una os dois pontos extremos da curva potencial, determinando-se os coeficientes da sua equação típica ($y = \alpha x + \beta$), em que α corresponde ao coeficiente angular da reta e “ β ” ao seu

coeficiente linear). Guarda-se a EQUAÇÃO 2, infradescrita, dirigindo-se ao terceiro passo.

$$y = \alpha x + \beta \text{ (EQUAÇÃO 2);}$$

- 3º Passo: Subtraindo-se a EQUAÇÃO 1 da EQUAÇÃO 2, obtém-se a EQUAÇÃO 3, abaixo:

$$y = (\beta + \alpha x) - (ax^{-b}) \text{ (EQUAÇÃO 3);}$$

- 4º Passo: Calcula-se a derivada primeira da EQUAÇÃO 3, obtendo-se a EQUAÇÃO 4. Igualando-se esta a zero e isolando-se a variável “x”, chega-se à expressão (EQUAÇÃO 5) da qual pode-se obter o valor da abscissa correspondente ao ponto de máxima curvatura (“X_{MCIII}”).

$$\frac{dy}{dx} = \alpha + (ba)x^{-b-1} \text{ (EQUAÇÃO 4);}$$

$$X_{MCIII} = (-\alpha / ba)^{(1/-b-1)} \text{ (EQUAÇÃO 5).}$$

Há vários métodos para a estimação dos parâmetros das EQUAÇÕES 1 e 2, cuja escolha seria de competência do pesquisador, tendo, Barros e Tavares (1995), utilizado um modelo matemático linear que levava em conta os graus de liberdade associados às observações dos CV, no tocante à EQUAÇÃO 1.

Apesar de já haver passado um certo tempo da publicação desta técnica (1995), ainda não se verificou nenhum trabalho que a tenha utilizado, no que se refere à determinação do tamanho de amostras (Oliveira (2011), avaliou-a no tocante ao dimensionamento de parcelas experimentais). É provável que isto se deva tanto à ampla divulgação dos métodos de Lessman e Atkins (1963), Meier e Lessman (1971) e Chaves (1985), quanto pelo desenvolvimento de novos modelos matemáticos, tal qual o “*relative CV model*” (modelo do CV relativo), desenvolvido recentemente por Carneiro *et al.*, (2012).

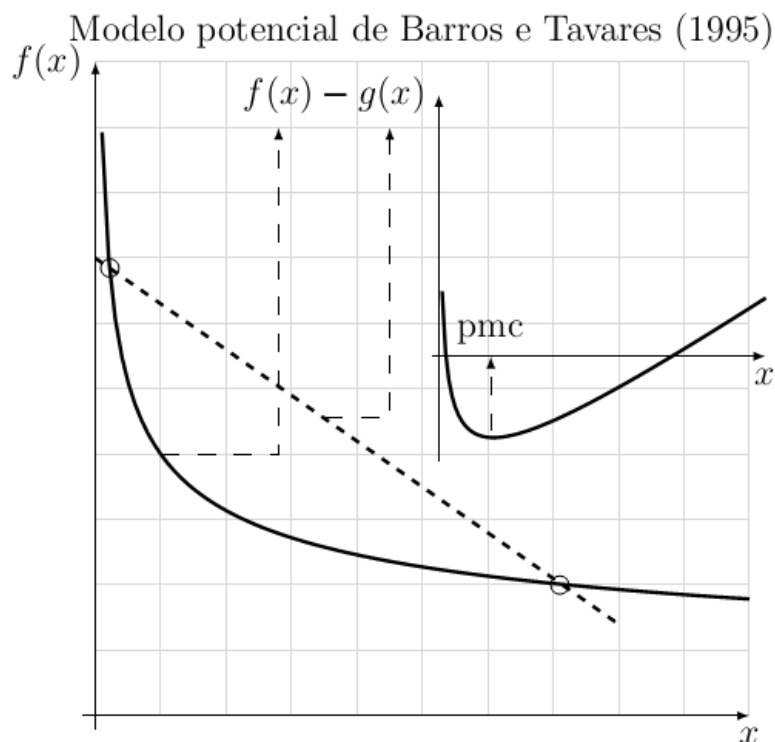


FIGURA 2: Esboço ilustrativo da obtenção do ponto de máxima curvatura da função potencial pelo método de Barros e Tavares (1995).

2.6.4 Modelos Segmentados com Resposta Platô

Diante das distintas abordagens matemáticas ao problema da determinação da máxima curvatura da curva potencial, muitos pesquisadores têm lançado mão de outros modelos não lineares que sejam capazes de estimar valores que representem as dimensões ideais de parcelas experimentais e amostras em parcelas. Mediante tais métodos, tem-se pretendido simplificar as interpretações gráficas e algébricas utilizadas para a determinação daqueles valores supracitados, oferecendo-se uma resposta singular à problemática em questão.

Segundo Ferreira (2011), modelos segmentados com resposta platô caracterizam-se por serem composições de dois modelos distintos, a saber: i) um modelo crescente ou decrescente até um determinado ponto crítico, cuja abscissa respectiva representaria a resposta procurada pelo pesquisador (a estimativa do parâmetro em discussão); e ii) um modelo linear constante (de coeficiente angular igual a zero) a partir do ponto citado anteriormente.

No caso específico do modelo linear segmentado com platô (FIGURA 3), a sua composição se daria pelo ajuste de um modelo não linear composto por duas retas, tendo como resultado um segmento de inclinação não nula até o ponto crítico e outro, subsequentemente, deste ponto em diante, o qual caracterizaria o platô. Ferreira (2011), descreve-o do seguinte modo:

$$f(x) = \begin{cases} \beta_0 + \beta_1 x, & x \leq x_0 \\ P = \beta_0 + \beta_1 x_0, & x > x_0 \end{cases},$$

no qual,

y: variável dependente;

β_0 e β_1 : parâmetros do segmento linear pré-platô;

x: variável independente;

x_0 : ponto crítico, de união entre o segmento pré-platô e o segmento platô;

P: valor da variável dependente no ponto crítico, responsável pela continuidade do modelo, sendo $P = \beta_0 + \beta_1 x_0$.

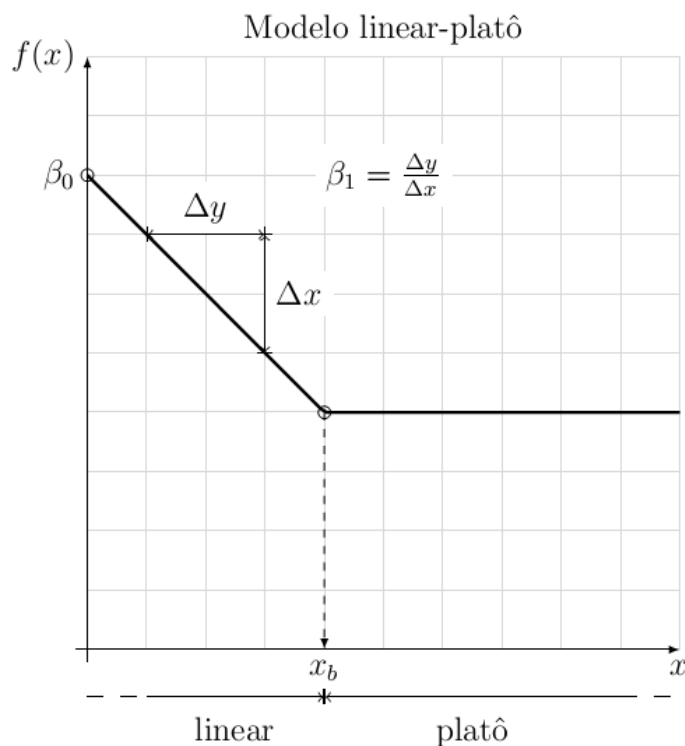


FIGURA 3: Esboço do modelo linear segmentado com platô.

Este modelo foi um dos utilizados por Moura (2008) na determinação do tamanho amostral para a avaliação de várias características dos frutos de híbridos de melão amarelo. Carneiro *et al.* (2012) utilizaram-no, conjuntamente com outros métodos, para a quantificação do número de plantas de feijoeiro necessárias para a avaliação da tolerância de campo à moléstia do mofo branco (*Sclerotinia sclerotiorum*).

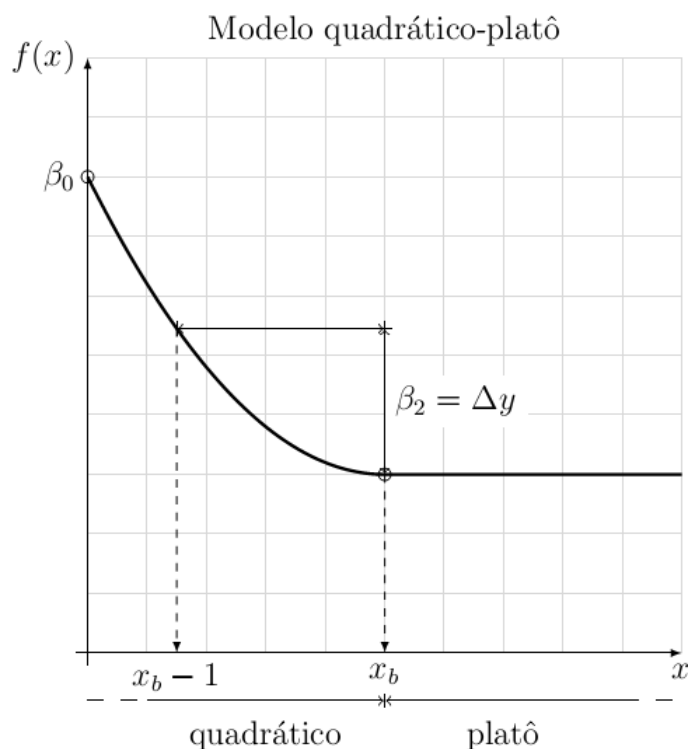


FIGURA 4: Esboço do modelo quadrático com resposta platô.

Quanto ao modelo quadrático com resposta platô (FIGURA 4), o seu ajuste resulta num segmento quadrático, o qual, no ponto crítico, une-se ao segmento linear platô. Conforme Ferreira (2011), este modelo pode ser representado por:

$$f(x) = \begin{cases} \beta_0 + \beta_1 x + \beta_2 x^2, & x \leq x_0 \\ P = \beta_0 - \frac{\beta_1^2}{4\beta_2}, & x > x_0 \end{cases}$$

no qual,

y: variável dependente;

β_0 , β_1 e β_2 : parâmetros do segmento linear pré-platô;

x: variável independente;

x_0 : ponto crítico, de união entre o segmento pré-platô e o segmento platô, sendo $x_0 = -\beta_1 / 2\beta_2$.

2.6.5 Modelo da Curva de Michaelis-Menten Reparametrizada ou Modelo do CV Relativo (“*Relative CV Model*”)

Trata-se de uma reparametrização do modelo de Michaelis e Menten⁴ (1913, citados por ZEVIANI, 2009), originalmente proposto para a explicação do comportamento de reações bioquímicas, incluindo-se nos modelos gerais de cálculos para a estimativa da meia vida de substâncias químicas das mais variadas (ZEVIANI, 2009).

Carneiro *et al.* (2012) desenvolveram este modelo (FIGURA 5) trabalhando na determinação do número mínimo de plantas de feijão comum – para a avaliação da resistência à *Sclerotinia sclerotiorum* (mofo branco) a campo, o qual, com ligeiras adaptações, pode ser descrito da seguinte forma:

$$CV = b_0 + \frac{b_1 - b_0}{1 + \frac{1-q}{q} \frac{N-1}{b_2 - 1}},$$

em que:

CV: valor do CV para um tamanho amostral composto por N unidades básicas;

b_0 : valor do CV para um tamanho amostral que tenda ao infinito;

b_1 : valor do CV para uma amostra unitária;

b_2 : tamanho amostral ideal para um valor pré-determinado de q; e

q: valor pré-determinado pelo pesquisador, relativo à amplitude máxima permitida para os valores de CV.

Talvez a principal vantagem deste modelo resida na possibilidade da sua resposta ser modulada pelo pesquisador, mediante a pré-determinação de um valor

⁴ MICHAELIS, L. e MENTEN, M. L. **Kinetics for Interface Action**. Biochemische Zeitung, Berlin, v.49, p.333–369, 1913.

percentual q , o qual varia de 0 a 1 (CARNEIRO *et al.*, 2012). Desta forma, para um valor nulo de q , o CV será mínimo e o tamanho da amostra tenderá ao infinito. Por outro lado, para um valor unitário de q , o CV será máximo e a amostra será composta por apenas uma unidade básica (FIGURA 4).

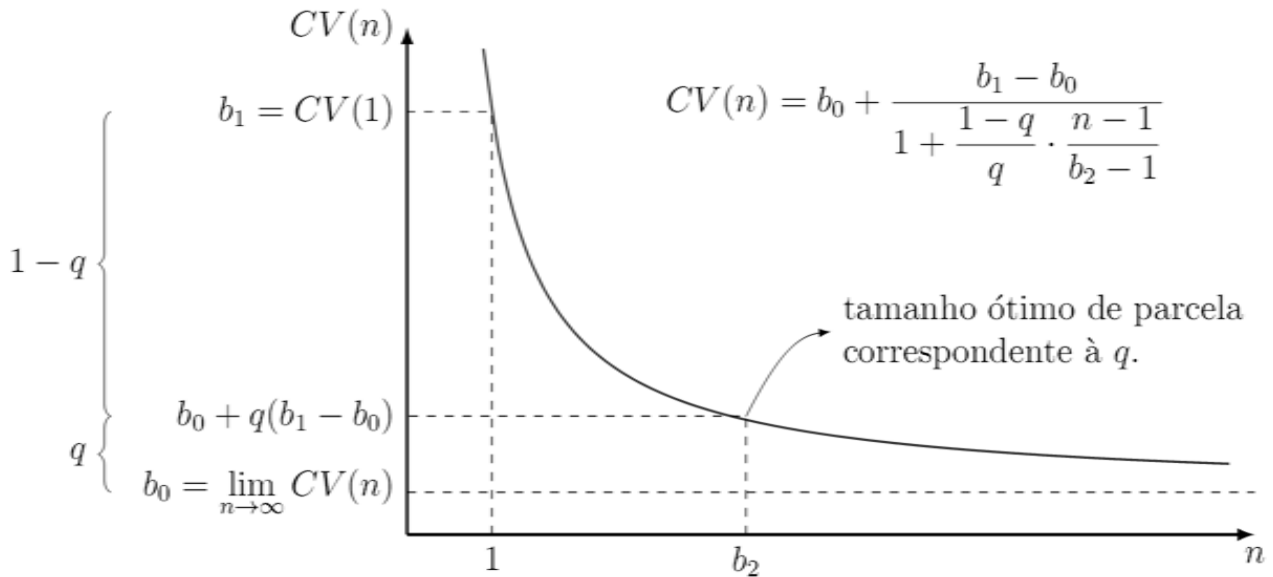


FIGURA 5: Gráfico explicativo do modelo do CV relativo, tal qual utilizado por Carneiro *et al.* (2012).

3. CAPÍTULO – SUFICIÊNCIA AMOSTRAL EM EXPERIMENTOS COM CANA-DE-AÇÚCAR

3.1 RESUMO

Para a avaliação de experimentos com cana-de-açúcar, tem-se recomendado a colheita total de parcelas, algo que, devido aos seus custos, tem-se mostrado inexecutável. Alternativamente, pode-se estimar a produtividade com base na tomada de amostras que representem as parcelas. O objetivo deste trabalho foi estudar a suficiência amostral em cana-de-açúcar para a estimação da massa e do teor de sólidos solúveis de parcelas experimentais. Também procurou-se definir o tamanho de amostra suficiente à avaliação experimental dos teores de sacarose da cultura. Reamostragens com reposição subsidiaram o uso dos modelos da curva potencial, linear e quadrático segmentados com resposta platô e do modelo do coeficiente de variação relativo. Também foram realizados estudos simulados de análise de variância. Os resultados indicaram que amostras de 20 colmos coletados aleatoriamente na parcela experimental permitem estimar adequadamente a variável massa do colmo. Com este valor, serão obtidas avaliações confiáveis da produtividade em colmos (TCH). Para o BRIX, leituras em 10 colmos de 10 touceiras aleatoriamente coletadas são suficientes à sua estimativa. Para as avaliações experimentais de sacarose (POL), podem ser usados feixes que contenham ao menos 20 colmos.

Palavras-chave: Cana-de-açúcar. Tamanho da amostra. Suficiência amostral. Melhoramento da cana-de-açúcar. *Saccharum* spp.

SAMPLE SUFFICIENCY IN SUGARCANE EXPERIMENTS

3.2 ABSTRACT

During the evaluation of the experimental phase, which corresponds to competition experiments and indication of cultivars, is recommended total harvest of plots, which, owing to their costs, can become impracticable. Optionally, one can estimate the productivity based on taking samples representing plots. The aim of this study was to research the sample sufficiency in sugarcane for estimating the plot mass and the brix contents of experimental plots. One tried to define also the sufficient sample size for the experimental evaluation of the sucrose contents. One used the maximum curvature of the power model, the linear and quadratic segmented models with response plateau and the relative CV model. Studies based on simulated analysis of variance were done also. The application of these models and the ANOVA was supported by bootstrap resamplings. The results indicated that a sample made by 20 stalks harvested in plot can be used for estimating properly the variable mass of the stalk. This value may be used for estimating the sugarcane yield per hectare. For brix variable, 10 stalks of 10 different plants are sufficient for estimating this variable. For POL, bundles containing at least 20 stalks can be sampled for estimating this variable. Samples with such dimensions can be used to estimate appropriately brix yield and sucrose yield per hectare.

Keywords: Sugarcane. Sample size. Sample sufficiency. Sugarcane breeding. *Saccharum* spp.

3.3 INTRODUÇÃO

A cana-de-açúcar (*Saccharum* spp.) encontra-se entre as mais importantes plantas cultivadas no mundo, ocupando, no Brasil, mais de oito milhões e meio de hectares sob cultivo (CONAB, 2012), o que contribui para que o país seja o líder global na produção e exportação de produtos da cadeia sucroenergética.

Devido ao ciclo produtivo e fenológico da cultura, as diferentes fases experimentais do melhoramento da cultura sobrepõem-se no tempo, gerando-se altíssimas demandas por mão-de-obra, insumos e áreas experimentais. Portanto, torna-se necessário que se otimize as operações envolvidas na condução das fases, especialmente aquelas relacionadas à fase experimental de competição de clones (FE), a qual é conduzida simultaneamente em diversos ambientes (estações experimentais). A colheita total das parcelas das centenas de ensaios em FE, ainda que desejável, têm-se tornado uma tarefa cada vez mais complicada e quase anti-econômica de se realizar. Como opção, podem ser efetuadas estimativas confiáveis de produtividade a partir de técnicas de amostragem que indiquem a suficiência amostral das variáveis que compõem o rendimento agrícola e industrial das culturas.

A possibilidade de se utilizar dados advindos de amostragem tem sido descrita em vários compêndios de Experimentação Agrícola, como no de Storck *et al.*, (2006). Contudo, ainda há lacunas referentes ao dimensionamento do tamanho de amostra ideal mediante estudos biométricos, especialmente no caso da cultura da cana-de-açúcar.

Qualquer amostra compõe-se de um determinado número de unidades básicas de amostragem, sendo evidente que com o aumento do número de unidades básicas amostradas, a precisão e a confiabilidade das estimativas calculadas também serão incrementadas (SILVA *et al.*, 1993). Porém, a partir de uma determinada dimensão, incrementos no número de unidades básicas amostradas não se traduzem em proporcionais aumentos de precisão, havendo pouca vantagem em se tomar amostras maiores (SILVA *et al.*, 1993). Portanto, a fim de se validar uma recomendação de suficiência amostral, deve-se determinar o ponto crítico de desaceleração das medidas de variação entre os diferentes tamanhos de amostra possíveis. Isto tem sido feito mediante metodologias que capazes de descrever a resposta desta variação a incrementos no tamanho da amostra.

Em substituição à colheita total de parcelas experimentais de cana-de-açúcar, Landell *et al.* (1999) avaliaram alguns métodos de amostragem para a estimativa do TCH em áreas sob colheita mecanizada, enquanto que Gheller *et al.* (1999) recomendaram a adoção de uma metodologia de amostragem para a avaliação indireta da produtividade, cujo enfoque concentrava-se mais nos procedimentos de amostragem em si do que no dimensionamento exato da amostra a ser coletada a campo. Neste sentido, o objetivo deste trabalho foi estudar a suficiência amostral para as estimativas das variáveis massa do colmo (MC), teor de sólidos solúveis (BRIX) e teor de sacarose (POL), que são as principais componentes estimáveis dos índices de produtividade em colmo, em sólidos solúveis e em sacarose da cultura da cana-de-açúcar, avaliando a sua aplicabilidade à estimação confiável destes três índices. A partir destes estudos, pretende-se estabelecer uma recomendação segura dos tamanhos de amostra a serem utilizados nos ensaios de competição de clones de cana-de-açúcar.

3.4 MATERIAIS E MÉTODOS

O experimento foi conduzido no âmbito do Programa de Melhoramento Genético da Cana-de-açúcar da Universidade Federal do Paraná, na subestação da cidade de Colorado, região centro-norte do estado do Paraná, cuja área situa-se às coordenadas 22° 45' 17,45'' S e 51° 56' 38,60'' W. O clima regional classifica-se como Cfa (subtropical úmido), com verões tépidos e chuvosos e invernos amenos. A área experimental encontra-se sobre Latossolo Vermelho distrófico de textura arenosa, num relevo “plano a suave ondulado”, a uma altitude de 400 metros.

O campo experimental, correspondente à FE, foi instalado no dia 18 de junho de 2010. O experimento consistia de 30 clones promissores oriundos de cruzamentos realizados na Estação de Florescimento e Cruzamento da Serra do Ouro, localizada no município de Murici, estado de Alagoas, no ano 2000 (série RB00). Destes clones, foram selecionados três para a elaboração do presente estudo. Adicionalmente avaliou-se a cultivar padrão RB867515. O delineamento utilizado foi o de blocos completos casualizados, com três repetições, tendo sido colhidas doze parcelas correspondentes aos três clones (RB005151, RB006655, RB006638) e à cultivar RB867515. Para as análises tecnológicas de sacarose, foram tomadas amostras das três parcelas da cultivar

padrão, RB867515, e do clone RB006613. As parcelas foram compostas de quatro linhas de 8,0 metros de comprimento, com espaçamento interlinear de 1,4 metros, resultando em uma área de 44,8 m². A densidade média de plantio foi de 18 gemas por metro. Este tamanho de parcela é compatível com as dimensões aceitáveis para a avaliação da massa MC, de acordo com Pontes *et al.* (2008).

A colheita das parcelas, cujas avaliações se deram em primeira cana soca, ocorreu na semana compreendida pelos dias 09 de julho a 13 de julho de 2012, tendo sido realizada da seguinte forma: (1) reconhecimento das parcelas e contagem (com os colmos eretos) do número de colmos industrializáveis de cada parcela; (2) leitura aleatória do BRIX de 3 colmos, a fim de que se pudesse antever o comportamento desta variável em cada parcela e em cada clone; (3) contagem e colheita de todos os colmos industrializáveis, separando-os conforme as suas respectivas plantas (touceiras); (4) descarte dos “ponteiros” (corte no ponto de quebra) e pesagem individual de cada colmo colhido, com balança de campo, cuja precisão nominal era de 10 gramas; (5) coleta do valor de BRIX de cada touceira, executada em um colmo industrializável por touceira, no quinto entrenó (da base do colmo ao seu ápice), usando-se instrumento coletor de caldo perfurante e refratômetro de campo, cuja precisão era de $\pm 0,2\%$; e (6) composição e amarração de feixes amostrais de 10, 20, 30 e 40 colmos para análises tecnológicas (determinação do teor percentual de sacarose do caldo – POL). Estes colmos foram coletados sequencialmente, a partir da primeira touceira da segunda linha, reiniciando-se a contagem após a obtenção de cada feixe. As unidades básicas de amostragem, conforme o exposto nos itens 3, 4, 5 e 6, acima, foram: o colmo, para a variável MC, touceira, para a variável BRIX, e feixes de colmos, para a variável POL.

Para a determinação da suficiência amostral, utilizaram-se as seguintes metodologias: (1) Máxima Curvatura do Modelo da Curva Potencial (LOGAN, 2010; MEIER e LESSMAN, 1971; e BARROS e TAVARES, 1995); (2) Modelo Linear Segmentado com Platô (FERREIRA, 2011); (3) Modelo Quadrático Segmentado com Platô (FERREIRA, 2011); e (4) Modelo do Coeficiente de Variação Relativo (CARNEIRO et al., 2012). Reamostragens *bootstrap* subsidiaram a utilização das metodologias supracitadas mediante a obtenção dos valores médios dos coeficientes de variação (CV) para cada tamanho de amostra, aos quais se ajustaram as curvas características desses modelos.

Os pontos de máxima curvatura do método da curva potencial, cujos valores de abscissa correspondem ao tamanho de amostra ideal, foram determinados pelas

seguintes proposições teóricas: (1) método da máxima curvatura modificado, ora abreviado como MCM I (MEIER e LESSMAN, 1971), usando-se a expressão: $x_{MCM I} = [a_0^2 \times (a_1^2(2a_1 + 1))/(a_1 + 2)]^{1/(2a_1+2)}$, em que $x_{MCM I}$ representa o ponto de máxima curvatura da função, e a_0 e a_1 são os parâmetros da curva potencial, estimados por regressão não linear; (2) método da máxima curvatura modificado de Barros e Tavares (1995), denominado MCM II. Este método se baseia na diferença entre a função da curva potencial ajustada e uma função linear cuja reta passe pelos pontos correspondentes ao menor e ao maior tamanho de amostra comparados, sendo estes dois pontos as intersecções entre ambas as funções. Desta maneira, a fórmula usada no cálculo para a determinação do ponto de máxima curvatura foi: $x_{MCM II} = (-a_3/a_1 a_0)^{1/(-a_1-1)}$, onde $x_{MCM II}$ representa o ponto de máxima curvatura da função, b_2 corresponde ao coeficiente angular da reta supramencionada, e b_0 e b_1 são os parâmetros da curva potencial, estimados por regressão não linear. O modelo de curva potencial utilizado foi $CV_i = a_0/n_i^{a_1}$ (LOGAN, 2010), em que n representa o i -ésimo tamanho de amostra.

Os modelos linear segmentado com resposta platô (LSP) e quadrático segmentado com resposta platô (QRP) foram ajustados aos conjuntos de dados obtidos de acordo com a sua descrição por Ferreira (2011), a saber: $CV_i = b_0 + b_1 n_i$, $n_i \leq n_0$ e $CV_i = b_0 + b_1 n_0$, $n_i > n_0$, onde b_0 e b_1 são os parâmetros do segmento linear pré-platô, n_i é o i -ésimo tamanho amostral, n_0 é o valor da abscissa no ponto de união entre o segmento linear pré-platô e o segmento platô, representando o tamanho da amostral ideal (LSP); e $CV_i = c_0 + c_1 n_i + c_2 n_i^2$, $n_i \leq n_0$ e $CV_i = c_0 - (c_1^2/4c_2)$, $n_i > n_0$, onde c_0 , c_1 e c_2 são os parâmetros do segmento quadrático pré-platô, n_i é o i -ésimo tamanho de amostra, n_0 é o valor da abscissa no ponto de união entre o segmento quadrático pré-platô e o segmento platô, representando o tamanho ideal de amostra (QRP).

No caso do modelo do CV relativo, a curva foi ajustada a partir do seguinte modelo teórico: $CV_i = d_0 + (d_1 - d_0)/(1 + ((1 - q)/q)((n_i - 1)/(d_2 - 1)))$, no qual CV_i é o valor do CV para uma amostra ideal de tamanho n_i , d_0 é o valor do CV para um tamanho amostral que tenda ao infinito, d_1 é o valor do CV de uma amostra unitária, d_2 representa o tamanho amostral ideal para um valor de q pré-determinado pelo pesquisador, relativo à amplitude máxima permitida para o valores da estatística CV. O diferencial deste modelo reside na possibilidade da sua resposta ser modulada pelo

pesquisador, a partir da pré-determinação de um valor percentual q , o qual varia de 0 a 1 (CARNEIRO *et al.*, 2012). Tendo-se adotado q igual 0,15, admitiu-se aceitar como ideal um tamanho de amostra cujo CV seja 15% maior do que o mínimo, que corresponde a um valor de q igual a zero, para o qual a amostra tenderia ao infinito.

Os ajustes dos modelos supramencionados foram subsidiados pela obtenção de dados através de reamostragens *bootstrap* (OYEYEMI, 2008). Foram executadas, para cada tamanho de amostra proposto, 1000 reamostragens com reposição, o que gerou 1000 reamostras distintas para cada tamanho amostral e para cada variável considerada. Estes procedimentos de amostragem aleatória foram realizados para cada uma das parcelas experimentais. Os tamanhos de amostra comparados foram os seguintes: 1, 2, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200 e 507 colmos (variável MC) e 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 e 53 touceiras (variável BRIX). Os valores extremos, 507 e 53, correspondem ao número de colmos e de touceiras das parcelas que apresentaram maior número de colmos e maior brotação, respectivamente.

No caso da variável POL, para cada tamanho de amostra, realizou-se uma análise de variância (ANOVA), de acordo com o seguinte modelo estatístico (STORCK *et al.*, 2006): $Y_{ij} = \mu + A_i + B_j + \varepsilon_{ij}$, no qual μ é a média geral de POL; A_i representa o efeito fixo do i -ésimo tamanho amostral, tendo variado de 1 a 4 (feixes de 10, 20, 30 e 40 colmos); B_j representa o efeito do j -ésimo bloco; e ε_{ij} é o erro experimental aleatório associado à observação Y_{ij} . Relacionando-se os CV experimentais oriundos das ANOVA aos tamanhos de amostra testados, ajustou-se o modelo da curva potencial ao conjunto de dados obtido. A indicação do tamanho de amostra suficiente à sua estimativa adveio da aplicação do método da máxima curvatura modificado por Barros e Tavares (1995), do mesmo modo como fôra utilizado no estudo das variáveis MC e BRIX.

A aplicabilidade da suficiência amostral obtida para a variável MC foi testada pela competição de clones, a fim de se avaliar a sua efetividade na estimação do índice de produtividade em colmos por hectare (TCH), calculado a partir dos dados biométricos de MC (média, no caso de dados amostrais, ou massa total de parcelas) e de número de colmos por metro (NC/M) através da seguinte expressão: $TCH = \overline{MC} \times NC/M^2 \times 10^{-2}$. Com base neste, no BRIX e no POL médios das parcelas, obtém-se

diretamente o índice de produtividade em sólidos solúveis ($TBH = (TCH \times \overline{BRIX})/100$) e em sacarose solúveis ($TPH = (TCH \times \overline{POL})/100$).

Para a avaliação do desempenho do tamanho médio de amostra indicado pelo estudo da variável MC na estimação do índice TCH, foram executadas 1000 análises de variância (ANOVA) simuladas para cada um dos tamanhos amostrais anteriormente testados. Como a análise executada com os dados obtidos da colheita total das parcelas dos quatro clones efetivamente colhidos a campo não havia detectado diferenças estatísticas entre tais clones, incluiu-se um clone testador aos demais, cujos dados virtuais foram simulados pelo Programa R. Tanto as médias dos blocos quanto a média geral deste clone testador foram projetadas para serem semelhantes à média nacional de produtividade em colmos. Além disto, a média deste testador diferiu dos demais clones para um nível de significância (α) igual a 0,05. Dos resultados destas 1000 análises e da análise executada com os dados obtidos da colheita total das parcelas (ANOVA real, chamada ANOVA-R), incluindo-se os dados simulados para o clone testador, foram extraídos os seguintes itens de comparação: (i) os p-valores dos testes F, a fim de se avaliar o poder do teste na detecção da diferença apontada pela ANOVA-R; e (ii) ranqueamento dos clones pelas suas médias de produtividade. Estes dados foram analisados e avaliados da seguinte forma: (i) análise dos p-valores obtidos pelos testes F, para verificar-se, proporcionalmente, o número de vezes em que resultados dos p-valores foram significantes ($\alpha \leq 0,05$), conforme o esperado; e (ii) comparação entre os *rankings* obtidos pela correlação de postos de Spearman.

Para estas ANOVA, adotou-se o seguinte modelo estatístico (STORCK *et al.*, 2006): $Y_{ijk} = \mu + G_i + B_j + \varepsilon_{ij} + a_{ijk}$, no qual μ é a média geral de produtividade; G_i representa o efeito fixo do i-ésimo genótipo testado, tendo variado de 1 a 5 (três clones, a cultivar padrão, mais o clone testador); B_j representa o efeito do j-ésimo bloco; ε_{ij} é o erro experimental aleatório associado à observação de produtividade em colmos Y_{ij} ; e a_{ijk} é o efeito aleatório do erro amostral.

Para a implementação de todos os procedimentos estatísticos utilizados ao longo deste trabalho, lançou-se mão do aplicativo R de computação estatística (R DEVELOPMENT CORE TEAM, 2012).

3.5 RESULTADOS E DISCUSSÃO

Os gráficos gerados pelos valores de CV obtidos das 1000 reamostragens (para cara amostral e em cada parcela) evidenciam o comportamento inversamente proporcional dos CV com relação ao tamanho amostral, seja para a variável MC (Figura 6), seja para a variável BRIX (Figura 7). Esta relação inversa também pode ser verificada no caso dos CV experimentais calculados no estudo da variável POL (Figura 8). A análise prévia destes gráficos demonstra a existência de um momento crítico de desaceleração (estabilização) dos CV, a partir do qual haveria pouca vantagem em se aumentar a quantidade de unidades básicas amostradas (SILVA *et al.*, 1993), sob pena de se aumentar excessivamente o tempo e os custos de amostragem para obter-se ganhos pouco expressivos na precisão experimental.

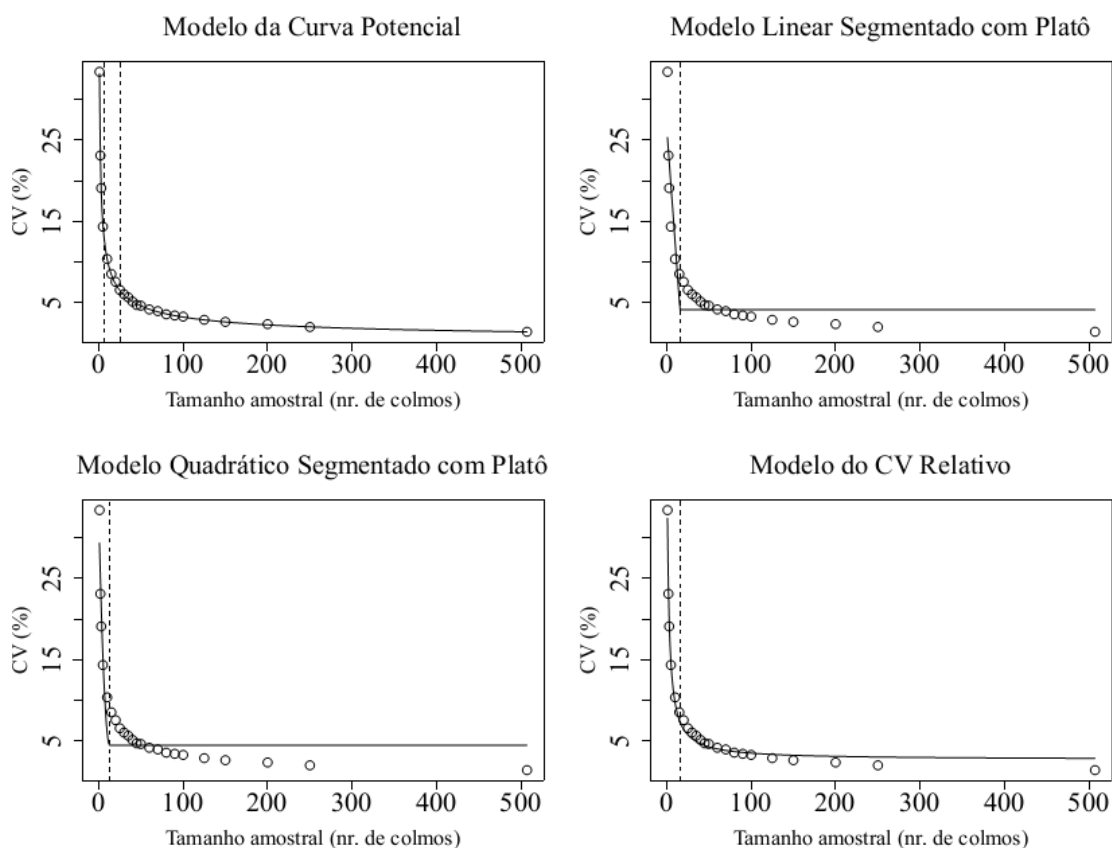


FIGURA 6: Gráficos dos ajustes dos modelos CP, LSP, QRP e CVR, para uma parcela (bloco um) da cultivar padrão. As linhas tracejadas remetem às abscissas dos pontos correspondentes à suficiência amostral obtida através de cada modelo, para a variável MC.

Os elevados valores de R^2 obtidos pelo ajuste do modelo da curva potencial (Tabela 1) assinalam a qualidade da sua adequação aos dados, fato igualmente observado por Araújo (2007) e por Pequeno *et al.* (2003), que obtiveram R^2 da ordem de 0,9686 e de 0,9800, respectivamente. Ademais, de acordo com o AIC (critério de informação de Akaike corrigido), dentre os modelos ajustados, este foi o que melhor explicou o comportamento dos CV de ambas as variáveis.

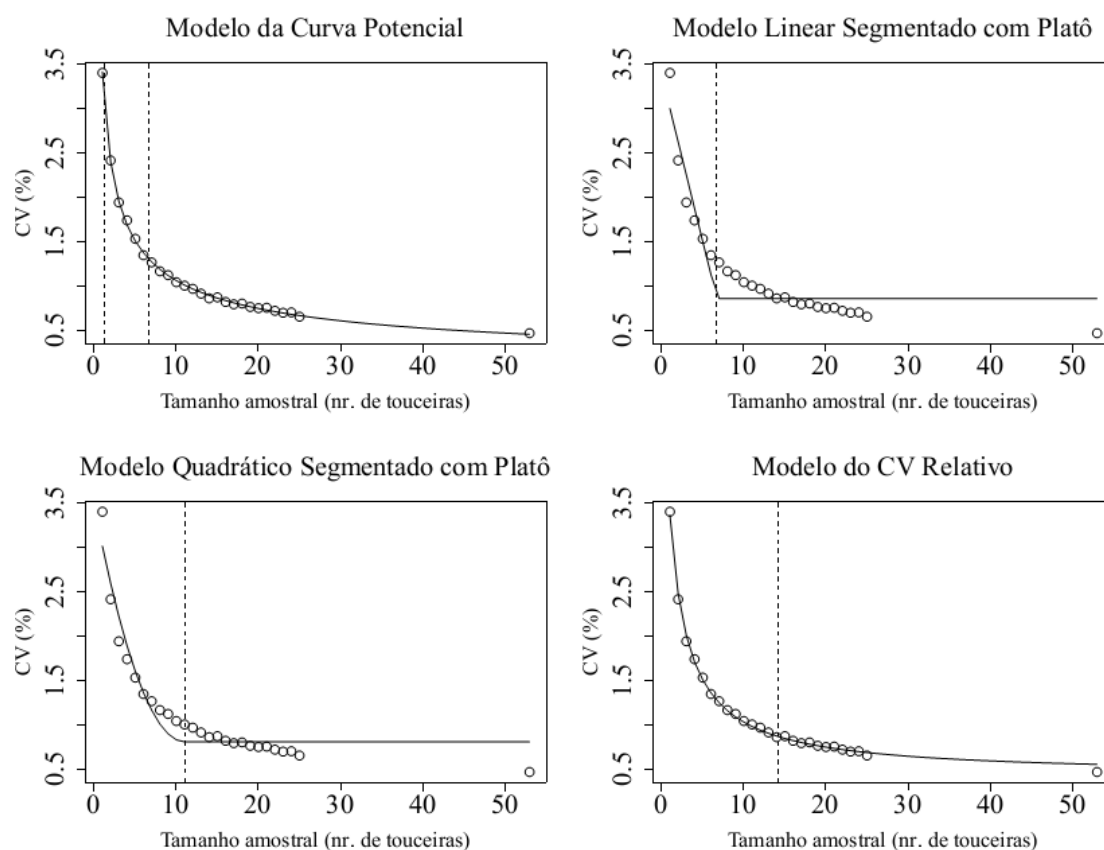


FIGURA 7: Gráficos dos ajustes dos modelos CP, LSP, QRP e CVR, para uma parcela (bloco um) da cultivar padrão. As linhas tracejadas remetem às abscissas dos pontos correspondentes à suficiência amostral obtida através de cada modelo, para a variável BRIX.

Aplicando-se o método MCM I à curva potencial gerada para cada parcela, obteve-se um tamanho médio de amostra (Tabela 2) de 7 colmos. Já pela aplicação do método MCM II, o valor médio obtido foi expressivamente maior: 35 colmos. Tais resultados corroboram a hipótese de Chaves (1985, citado por OLIVEIRA *et al.*, 2011), para quem o método da MCM I subestimaria os valores de abscissa dos pontos de máxima curvatura, representando muito mais um tamanho mínimo da amostra do que o ideal. Portanto, apesar do método da MCM I ter apresentado resultados satisfatórios para Araújo *et al.*, (2009), considerando-se que nos experimentos em FE apenas clones

de ótima performance (clones elites) são postos em competição, havendo, no geral, sutis diferenças a serem detectadas entre eles, seria prudente basear-se em outros métodos para a determinação do tamanho ideal de amostra para a avaliação da variável MC. Neste sentido, para o dimensionamento da amostra a ser utilizada na seleção de famílias da fase T1 do melhoramento da cultura da cana-de-açúcar, o MCM I foi criticado por Leite (2007).

TABELA 1: Valores de R^2 e de AIC (Critério de Informação de Akaike) para os modelos CP, LSP, QRP e CVR ajustados às variáveis MC e BRIX.

MC								
Clone / Bloco	R^2				AIC			
	CP	LSP	QRP	CVR	CP	LSP	QRP	CVR
RB867515 / 1	0,9997	0,8770	0,9158	0,9903	-19,5617	120,8377	112,1191	62,3454
RB867515 / 2	0,9995	0,8940	0,9252	0,9933	-16,0275	111,9488	103,9472	48,5917
RB867515 / 3	0,9996	0,8776	0,9146	0,9902	-17,5065	115,8417	107,5569	57,8811
RB005151 / 1	0,9995	0,8904	0,9253	0,9926	-10,3770	117,8174	108,9971	55,7377
RB005151 / 2	0,9993	0,8778	0,9132	0,9894	-10,7858	112,0763	104,2175	55,8948
RB005151 / 3	0,9988	0,8665	0,9207	0,9902	4,5690	116,6343	104,6648	56,4895
RB006655 / 1	0,9998	0,8858	0,9219	0,9921	-26,7848	122,7890	114,0354	61,4042
RB006655 / 2	0,9991	0,8775	0,9141	0,9891	1,1512	116,2204	108,0451	60,5237
RB006655 / 3	0,9993	0,8843	0,9196	0,9908	-15,3023	106,4323	98,0665	48,2521
RB006638 / 1	0,9996	0,8755	0,9167	0,9906	0,9996	115,3958	106,1494	55,8872
RB006638 / 2	0,9991	0,8833	0,9189	0,9901	4,2426	120,2621	111,9028	63,5547
RB006638 / 3	0,9991	0,8908	0,9183	0,9914	7,7763	121,3941	114,7035	63,0390
BRIX								
Clone / Bloco	R^2				AIC			
	CP	LSP	QRP	CVR	CP	LSP	QRP	CVR
RB867515 / 1	0,9991	0,8951	0,9335	0,9976	-126,0232	0,8247	-11,0032	-96,9104
RB867515 / 2	0,9985	0,8769	0,9251	0,9952	-100,1259	17,0860	4,1653	-67,0955
RB867515 / 3	0,9985	0,8738	0,9273	0,9964	-112,6297	5,9942	-8,3420	-86,1913
RB005151 / 1	0,9981	0,8693	0,9288	0,9961	-108,3126	4,1789	-11,6147	-87,3021
RB005151 / 2	0,9978	0,8706	0,9276	0,9961	-107,1308	1,6821	-13,4345	-89,2347
RB005151 / 3	0,9971	0,8784	0,9321	0,9960	-86,3366	13,8768	-1,2775	-74,9390
RB006655 / 1	0,9984	0,8809	0,9246	0,9947	-88,2749	26,3467	14,4690	-54,7014
RB006655 / 2	0,9980	0,8758	0,9300	0,9961	-104,7593	5,9982	-8,9024	-84,0931
RB006655 / 3	0,9988	0,8683	0,9224	0,9954	-118,3498	7,1350	-6,6204	-80,3429
RB006638 / 1	0,9984	0,8693	0,9170	0,9946	-104,1831	13,1961	1,3973	-69,8118
RB006638 / 2	0,9984	0,8651	0,9126	0,9942	-102,1678	16,4072	5,1162	-65,1982
RB006638 / 3	0,9986	0,8925	0,9295	0,9966	-94,4817	22,0043	11,0566	-67,7048

CP, modelo da curva potencial; LSP, modelo linear segmentado com platô; QRP, modelo quadrático segmentado com platô; e CVR, modelo do CV relativo.

Nota: Quanto menor o AIC, melhor será o ajustamento de um modelo aos dados, em comparação com os demais modelos ajustados.

A seu turno, o método da MCM II, desenvolvido por Barros e Tavares (1995) e até o momento sem registros de aplicação formal nos estudos de suficiência amostral, propiciou respostas bastante condizentes com o que se tem verificado empiricamente. O valor médio obtido através deste método aproxima-se tanto dos tamanhos práticos de amostra recomendados por Gheller *et al.* (1999) aos programas brasileiros de

melhoramento genético da cana-de-açúcar, quanto do número sugerido por Mariotti e Lascano (1969, citados por ARIZONO *et al.*, 1998), para quem 30 colmos ofereceriam boas estimativas da massa total da parcela.

Considerando-se as sugestões de Gheller *et al.* (1999) e o que empiricamente tem sido verificado a campo, o método CVR também poderia ser bem avaliado. Diante dos demais, resultados, existe a possibilidade de que os métodos QRP e LSP tenham subestimado sutilmente o tamanho de amostra ideal. Quanto ao LSP, contudo, Paranaíba *et al.* (2009) chegaram a resultados satisfatórios, utilizando esta metodologia.

TABELA 2: Dimensões da amostra ideal para as variáveis MC e BRIX, segundo os métodos MCM I, MCM II, LSP, QRP e CVR. Todos os resultados foram arredondados para o maior valor inteiro subsequente.

Clone / Bloco	CV%	MC (n° de colmos)					CV%	BRIX (n° de touceiras)				
		MCM I	MCM II	LSP	QRP	CVR (q=0,15)		MCM I	MCM II	LSP	QRP	CVR (q=0,15)
RB867515 / 1	33,02	7	27	16	14	17	3,39	2	7	7	12	15
RB867515 / 2	30,12	6	34	16	14	17	3,41	2	8	6	12	15
RB867515 / 3	37,19	7	34	16	18	20	5,08	2	9	7	12	15
RB005151 / 1	25,41	6	38	16	14	17	3,99	2	10	6	11	14
RB005151 / 2	27,77	6	39	16	14	18	3,27	2	9	6	12	16
RB005151 / 3	25,53	5	42	16	14	19	3,49	2	8	6	12	15
RB006655 / 1	29,80	6	36	16	14	17	3,48	2	9	6	12	16
RB006655 / 2	35,38	7	34	12	14	18	5,31	2	9	7	12	15
RB006655 / 3	33,90	7	31	16	14	17	4,14	2	7	6	11	14
RB006638 / 1	29,79	6	33	16	15	19	4,51	2	9	7	12	15
RB006638 / 2	28,64	6	36	16	10	15	4,06	2	9	6	12	15
RB006638 / 3	29,02	6	35	16	13	17	3,81	2	9	6	11	13
Média	30,46	7	35	16	14	18	3,99	2	9	7	12	15
Média geral	--			18			--			9		
Indicação prática	--			20			--			10		

MCM I, método da máxima curvatura da curva potencial I; MCM II, método da máxima curvatura da curva potencial II; LSP, modelo linear segmentado com platô; QRP, modelo quadrático segmentado com platô; e CVR, modelo do CV relativo.

Nota: -- As médias gerais referem-se às médias dos resultados obtidos pelos métodos, não se aplicando à coluna dos CV, os quais indicam a variação original existente em cada parcela, considerando-se a totalidade das unidades básicas.

Com relação aos ajustes para a variável BRIX, deve-se, em primeiro lugar, atentar para a pequena variação existente em cada parcela, sendo os seus CV, em média, quase oito vezes menores do que aqueles verificados para a variável MC, o que se trata de algo esperado para um caractere de alta herdabilidade expresso em populações

clonais (SILVA *et al.*, 2011), portanto menos sujeito a variações ambientais. De fato, ao proporem uma reclassificação para os CV em cana-de-açúcar, Couto *et al.* (2009) constataram que o BRIX apresentava a menor variabilidade dentre oito variáveis estudadas. Para esta variável, à exceção do MCM I, todos os demais métodos testados forneceram resultados semelhantes. No entanto, o método da MCM II e os modelos QRP e CVR parecem ter apresentado algum nível de superestimação (Tabela 2). De fato, os valores obtidos pela aplicação destes modelos aos dados ultrapassam aquele sugerido por Barbosa *et al.* (2001) para a estimação do BRIX de parcelas compostas por famílias de cana-de-açúcar, aproximando-se também aos tamanhos de amostra indicados por Leite *et al.* (2009), para avaliações do BRIX na fase T1. Nesta fase espera-se observar maior variabilidade em BRIX nas parcelas do que em FE, cujas populações são clonais. Com respeito ao método da MCM I, Barbosa *et al.* (1999) obtiveram bons resultados ao utilizá-lo na estimativa da suficiência amostral para avaliações do BRIX. Estes autores trabalharam na determinação do tamanho de amostra de famílias de irmãos germanos em experimentos da fase T1 e chegaram a valores condizentes com aqueles apresentados nas suas citações bibliográficas. Neste trabalho, contudo, este método pode ter apresentado indicações subestimadas, considerando-se a sua performance no estudo relativo à variável MC.

Oliveira (2011), na determinação do tamanho de parcelas de *Eremanthus erythropappus* (candeia), afirmou que o método da MCM II poderia apresentar insensibilidade à natureza dos caracteres sob avaliação, oferecendo resultados semelhantes para variáveis cujos comportamentos sejam bastante diferentes. Neste trabalho, porém, tal fato não se verificou.

Os trabalhos referentes à variável POL tampouco têm detectado expressivas variações experimentais, mesmo em estudos de seleção de famílias, nos quais se poderia esperar a obtenção de CV experimentais mais altos (SOUZA *et al.*, 2011; e DUTRA FILHO *et al.*, 2011). Couto *et al.* (2009) constataram que o POL posicionava-se em segundo lugar, em termos de variabilidade experimental, dentre oito variáveis estudadas. No presente trabalho, todos os CV experimentais estimados pelas ANOVA (Tabela 3), independente das dimensões dos feixes amostrados, posicionaram-se dentro do intervalo constituído pelos valores mínimo e máximo de 166 CV experimentais encontrados por Couto *et al.* (2013) na literatura especializada. Esta compilação serviu de base a estes autores para proposição de uma nova classificação para os CV de experimentos com cana-de-açúcar. Segundo esta nova classificação, todos os CV

experimentais obtidos enquadraram-se nas classes “médio” ou “baixo”, tendo ficado abaixo dos 10%. Para Gomes (2009), experimentos de campo com CV de até 10% seriam de boa precisão.

TABELA 3: CV experimentais – variável POL – obtidos pelas ANOVA executadas para cada tamanho de amostra e comparações das suas estimativas aos valores relatados por Couto *et al.* (2013).

	Tamanho do feixe (nr. de colmos)				Estimativas dos 4 experimentos		Estimativas de 166 experimentos	
	10	20	30	40	Média	D-P	Média	D-P
CV exp. (%)	5,85	4,45	2,88	1,85	3,76	1,75	6,46	3,97

CV exp., coeficiente de variação experimental; D-P, desvio-padrão.

Aplicando-se o MCM II à curva potencial ajustada aos dados experimentais de CV pelas feixes de diferentes tamanhos, obteve-se um valor correspondente a um feixe composto por 20,51 colmos (Figura 8), ou seja, vinte colmos e meio, aproximadamente. Este valor está em concordância com as recomendações práticas feitas por Gheller *et al.* (1999) aos programas de melhoramento genético de cana-de-açúcar. Estes autores sugerem que se envie um feixe que contenha a partir de 10 (10, 12, 15, etc.) colmos às análises tecnológicas, por parcela. Em termos de amostragem, poderia optar-se pela amarração de um feixe de 20 ou 21 colmos. O arredondamento pelo menor valor, neste caso, pode ser indiretamente sustentado pela análise do comportamento desta variável a campo, a partir da análise dos dados coletados por Couto *et al.* (2013). De fato, o valor de CV experimental de 4,45% obtido na análise efetuada com feixes de 20 colmos encontra-se abaixo da média relatada por aqueles autores. Ademais, os valores de desvio-padrão indicam ter havido menos variação nas análises realizadas neste trabalho do que naqueles 166 experimentos avaliados por Couto *et al.* (2013).

Dos resultados das 1000 ANOVA simuladas, nota-se que amostras de tamanho igual a 20 colmos, quando usadas para o cálculo do índice de produtividade em colmos, geraram análises com razoável poder de detecção de uma determinada diferença estatística entre os índices TCH dos clones em competição. De fato, efetuando-se simulações de ANOVA com este tamanho amostral, mais de 75% dos p-valores vinculados aos testes F das simulações detectaram a diferença existente (p-valor igual a 0,011) entre o clone testador e os demais clones, ou seja, foram iguais ou menores a 0,05 (Figura 9). Não foram localizados resultados, em cana-de-açúcar, aos quais estes pudessem ser diretamente comparados. No entanto, na avaliação de tamanhos de amostra e métodos de amostragem em milho, Silva *et al.* (1998) encontraram diferenças

nos resultados da interpretação da estatística F (dos p-valores, por conseguinte) com o aumento do tamanho da amostra, para os caracteres de comprimento e peso da espiga.

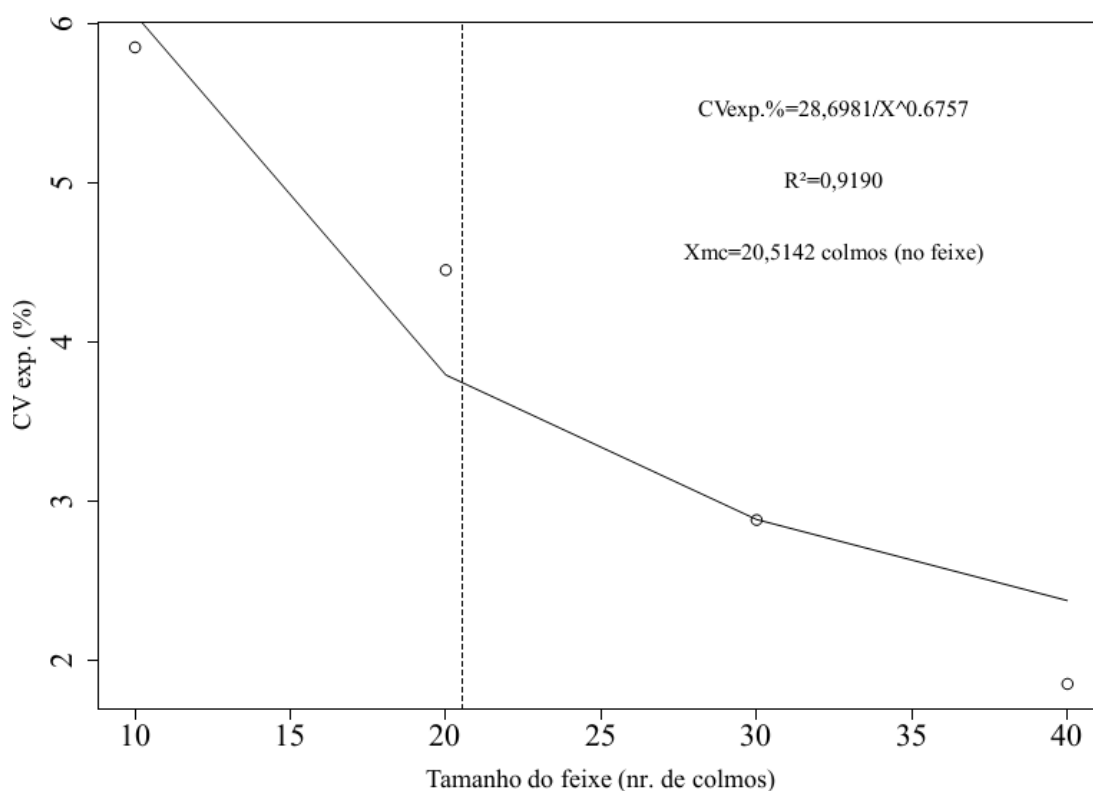


FIGURA 8: Gráfico do ajuste do modelo CP aos dados de CV exp. da variável POL em função do tamanho da amostra. A linha tracejada remete à abscissa do ponto correspondente à suficiência amostral obtida pelo MCM II.

Havendo ou não diferenças estatísticas detectáveis entre os clones, pode-se considerar o simples ranqueamento das médias entre os clones em competição como sendo uma estratégia auxiliar às análises experimentais, sobre as quais se fundamentará a seleção dos melhores genótipos nesta fase do melhoramento da cultura. Desta forma, comparando-se os *rankings* obtidos das 1000 ANOVA simuladas com o *ranking* oriundo da ANOVA-R, através da correlação de postos de Spearman (ρ), mais de 70% dos *rankings* gerados, no caso das amostras compostas por 20 colmos, forneceram resultados com apenas uma alteração de ordem (valores de ρ iguais ou maiores do que 0,9).

A importância de uma correta estimativa de TCH para a posterior estimativa dos índices derivados TBH e TPH pode ser observada em resultados de estudos recentes, como os de Dutra Filho *et al.* (2011) e de Souza *et al.* (2011), nos quais foram encontrados altíssimos valores de correlação entre TCH e TPH, da ordem de 98,0% e de

98,3%, respectivamente. No mesmo trabalho de Dutra Filho *et al.*, valores de correlação simples entre TPH e POL de apenas 0,26 e entre TCH e BRIX negativos (-0,16) reforçam a constatação de que as variáveis TPH e TBH dependeriam muito mais da produtividade em órgãos armazenadores, os colmos, do que dos teores das substâncias neles armazenadas, seja sacarose, sejam quaisquer outros sólidos solúveis.

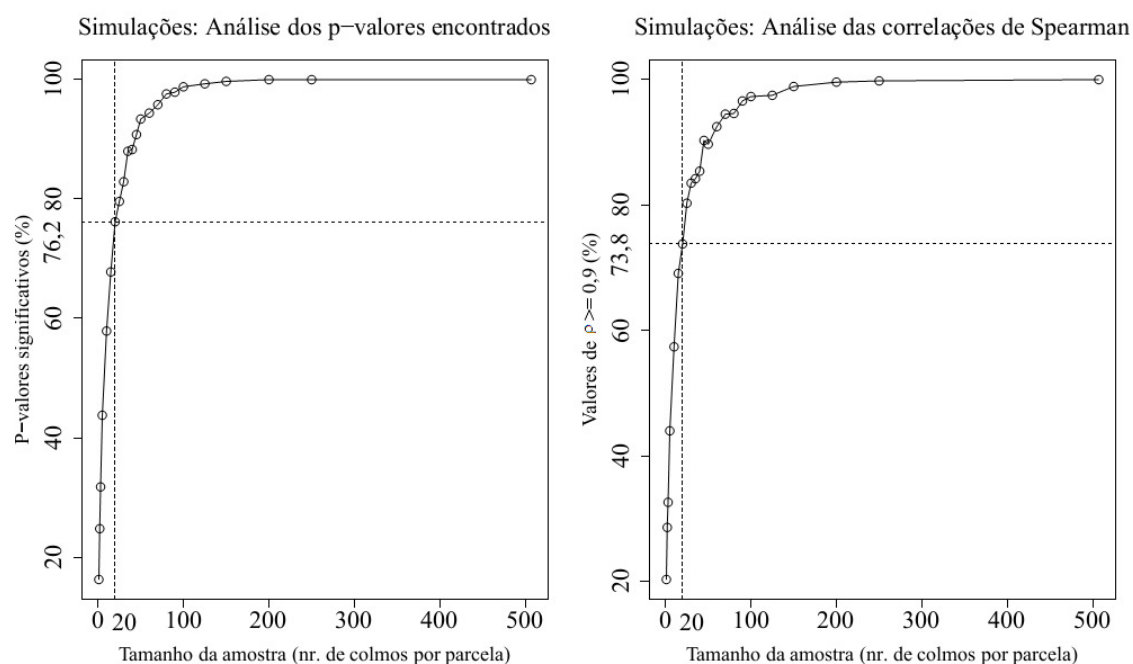


FIGURA 9: Para cada tamanho de amostra, proporções dos p-valores significativos e dos *rankings* cujos valores da correlação de Spearman foram maiores do que 0,9.

Os tamanhos de amostra indicados por este trabalho como sendo suficientes para os cálculos das estimativas de MC, BRIX e POL, e aplicáveis à composição dos índices de produtividade TCH, TBH e TPH, podem ser, pelas suas dimensões, aplicados às rotinas de amostragem dos programas de melhoramento genético de cana-de-açúcar, otimizando-se a utilização dos recursos destinados às avaliações e abreviando-se o tempo necessário à obtenção dos dados advindos do campo.

3.6 CONCLUSÕES

1. A massa média de vinte colmos industrializáveis, aleatoriamente coletados, é suficiente para a estimação da massa de parcelas experimentais de cana-de-açúcar.
2. O BRIX médio de dez touceiras aleatoriamente escolhidas, obtido de um colmo industrializável de cada uma delas, é suficiente para a estimação do teor de sólidos solúveis de parcelas experimentais de cana-de-açúcar.
3. Um feixe composto por 20 colmos industrializáveis, coletados sequencialmente a partir do início do segundo sulco de parcelas experimentais de cana-de-açúcar, é suficiente para que se estime o POL (teor de sacarose) de parcelas experimentais de cana-de-açúcar, a partir da análise tecnológica laboratorial do seu caldo.

3.7 CONSIDERAÇÕES FINAIS

Diante das conclusões obtidas por meio deste trabalho, pode-se sugerir a seguinte fórmula para o cálculo do índice de produtividade em colmos da cultura da cana-de-açúcar (TCH): $TCH = \overline{M20} \times NC/M^2 \times 10^{-2}$, em que $\overline{M20}$ é a massa média (g) de vinte colmos industrializáveis aleatoriamente coletados na parcela experimental e NC/M^2 é o número de colmos por área da parcela (m^2).

Aliando-se as estimativas de TCH calculadas mediante a fórmula sugerida acima ao valor prático de 10 touceiras para estimar-se o BRIX médio das parcelas, poderá ser adequadamente estimado o índice derivado TBH (produtividade em BRIX da cana-de-açúcar), através da seguinte expressão: $TBH = (TCH \times \overline{B10})/100$, na qual $\overline{B10}$ é o BRIX médio de dez touceiras aleatoriamente escolhidas. Por semelhante modo, unindo-se as estimativas de TCH às de POL, poderá ser adequadamente estimado o índice derivado TPH (produtividade em BRIX da cana-de-açúcar), através da seguinte expressão: $TPH = (TCH \times \overline{F20})/100$, na qual $\overline{F20}$ é o POL obtido da análise de um feixe de 20 colmos, coletados sequencialmente a partir do início do segundo sulco de parcelas experimentais.

3.8 REFERÊNCIAS

ARIZONO, H.; MATSUOKA, S.; GHELLER, A. C. A.; MATSUDA, Y.; HOFFMANN, H. P.; BASSINELLO, A. I.; MANESES, L. L. **Alternativas para avaliação de produção de cana-de-açúcar.** Stab, v.16, n.5, p.20, 1998.

ARAÚJO, E. G.; PIEDADE, S. M. DE S.; CONCEIÇÃO, M. A. F.; PEREIRA, J. C. **Métodos de amostragem e tamanho de amostra para avaliar o estado de maturação da uva niágara rosada.** Rev. Bras. Biom., São Paulo, v.27, n.4, p.501-518, 2009.

BARBOSA, M. H. P.; SILVEIRA, L. C. I.; DONIZETTE, R. **Tamanho da amostra para avaliação de famílias de cana-de-açúcar.** Anais do 7º Congresso STAB, p.57-59, 1999.

BARBOSA, M. H. P.; PETERNELLI, L. A.; SILVEIRA, L. C. I. **Plot size in sugarcane family selection experiments.** Crop Breeding and Applied Biotechnology, v. 1, n. 3, p. 271-276, 2001.

BARROS, I. e TAVARES, M. **Estimativa do tamanho ótimo de parcelas experimentais através de cálculos algébricos.** Bragantia, Campinas, 54(1): 209-215, 1995.

CARNEIRO, F. F.; ZEVIANI, W. M.; SANTOS, J. B.; CARVALHO, R. S. B.; ALVES, F. C.; DIAS, J. A. **Minimum number of common bean plants per plot to assess field resistance to white mold.** Crop Breeding and Applied Biotechnology 11: 358-364, 2011.

CONAB - Companhia Nacional de Abastecimento. **Acompanhamento de safra brasileira: cana-de-açúcar, terceiro levantamento, dezembro/2012 - companhia nacional de abastecimento.** – Brasília : Conab 2012.

COUTO, M. F.; PETERNELLI, L. A.; BARBOSA, M. H. P. **Proposta de classificação dos coeficientes de variação para a cultura da cana-de-açúcar.** In: 54^a. Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria *et* 13^o Simpósio de Estatística Aplicada à Experimentação Agronômica, 2009.

COUTO, M. F.; PETERNELLI, L. A.; BARBOSA, M. H. P. **Classification of the coefficients of variation for sugarcane crops.** Ciência Rural, Santa Maria, v.43, n.6, jun, 2013.

DUTRA FILHO, J. A.; MELO, L. J. O. T.; SIMÕES NETO, D. E.; ANUNCIAÇÃO FILHO, C. J.; BASTOS, G. Q.; DAROS, E. **Seleção de progênies e correlação de componentes de produção em cana-de-açúcar.** Revista Brasileira de Ciências Agrárias, Recife, v.6, n.3, p.432-439, 2011.

FERREIRA, D. F. **Recursos computacionais utilizando R.** Universidade Federal de Lavras, Departamento de Ciências Exatas. 2011.

GHELLER, A. C. A.; MENEZES, L. L.; MATSUOKA, S.; MASUDA, Y.; HOFFMANN, H. P.; ARIZONO, H.; GARCIA, A. A. F. **Manual de método alternativo para medição da produção de cana-de-açúcar.** Araras: UFSCar – CCA – DBV, 1999.

LANDELL, M. G. A.; VASCONCELOS, A. C. M.; SILVA, M. A.; PERECIN, D.; CARVALHO, R. S. R.; BARBOSA, V.; PENNA, M. J. **Validação de métodos de amostragem para a estimativa da produção de cana-de-açúcar, em áreas de colheita mecanizada.** STAB – novembro/dezembro/1999 – Vol. 18 n^o 2, p.48-51.

LEITE, M. S. O. **Tamanho da amostra para seleção de famílias de cana-de-açúcar.** Dissertação (mestrado). Universidade Federal de Viçosa, Viçosa, MG, 2007.

LEITE, M. S. O.; PETERNELLI, L. A.; BARBOSA, M. H. P.; CECON, P. R.; CRUZ, C. D. **Sample size for full-sib family evaluation in sugarcane.** Pesquisa agropecuária brasileira, Brasília, v.44, n.12, p.1562-1574, dez, 2009.

LOGAN, M. **Biostatistical design and analysis using R – a practical guide**. 1st. Ed. Wiley-Blackwell. Oxford, UK. 2010.

MEIER, V.D. e LESSMAN, K.J. **Estimation of optimum field plot shape and size testing yield in *Crambe abyssinica* HORDNT**. Crop Science, v.11, p.648-650, 1971.

OLIVEIRA, G. M. V. **Tamanho de parcelas experimentais para *Eremanthus erythropappus***. Dissertação (mestrado). Universidade Federal de Lavras, Lavras, MG, 2011.

OLIVEIRA, G. M. V.; MELLO, J. M.; LIMA, R. R.; SCOLFORO, J. R. S.; OLIVEIRA, A. D. **Tamanho e forma de parcelas experimentais para *Eremanthus erythropappus***. Cerne, Lavras, v. 17, n. 3, p. 327-338, jul./set. 2011.

OYEYEMI, G. M. **Comparison of bootstrap and jackknife methods of resampling in estimating population parameters**. Global Journal of Pure and Applied Sciences, Vol 14, No. 2, 2008: 217-220.

PARANAÍBA, P. F.; FERREIRA, D. F.; MORAIS, A. R. **Tamanho ótimo de parcelas experimentais: proposição de métodos de estimação**. Rev. Bras. Biom., São Paulo, v.27, n.2, p.255-268, 2009.

PEQUENO, S. A.; PINHEIRO, J. B.; ZUCCHI, M. I.; VENCONVSKY, R.; COELHO, A. S. G.; TRINDADE, M. G. **Determinação do número de marcadores RAPD para estudos da diversidade genética em soja utilizando o método *bootstrap***. Biosci. J., Uberlândia, v. 19, n. 2, p. 45-48, May/Aug. 2003.

PONTES, M. S.; PETERNELLI, L. A.; CARVALHO, M. P.; ANONI, C. O.; PEREIRA, R. C.; COUTO, M. F. **Estimativa do tamanho de parcela em cana-de-açúcar pelo método da máxima curvatura modificado**. Anais do 9º Congresso Nacional da Stab. Maceió, AL, 16 a 21 de novembro de 2008, p.636-640.

R CORE DEVELOPMENT TEAM **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2012, <http://www.R-project.org/>.

SILVA, G. C.; OLIVEIRA, F. J.; ANUNCIAÇÃO FILHO, C. J.; SIMÕES NETO, D. E.; MELO, L. J. O. T. **Divergência genética entre genótipos de cana-de-açúcar**. Revista Brasileira de Ciências Agrárias. Recife, v.6, n.1, p.52-58, 2011.

SILVA, J.; LIMA E SILVA, P. S.; NUNES, R. P. **Amostragem e tamanho da amostra na estimação de caracteres da espiga do milho**. Pesquisa agropecuária brasileira, Brasília, v.33, n.12, p.1983-1988, dez. 1998.

SILVA, P. S. L.; SILVA, G. S.; RODRIGUES, M. V.; MACHADO, A. A. **Métodos de amostragem e tamanho da amostra para alguns caracteres do milho**. Ciên. Agron., Fortaleza, 24 (1/2), pág 5-10, Junho/Dezembro. 1993

SOUZA, A. E. R.; OLIVEIRA, F. J.; ANUNCIAÇÃO FILHO, C. J.; MELO FILHO, P. A.; MELO, L. J. O. T.; TABOSA, J. N. **Seleção de famílias RB visando à alta produtividade e precocidade na maturação em cana-de-açúcar**. Bragantia, Campinas, v. 70, n. 4, p.788-795, 2011.

STORCK, L. GARCIA, D. C., LOPES, S. J., ESTEFANEL, V. **Experimentação vegetal**. Segunda Edição. Editora da UFSM. Santa Maria, RS, 2006.

4. REFERÊNCIAS GERAIS

ARAÚJO, E. G.; PIEDADE, S. M. DE S.; CONCEIÇÃO, M. A. F.; PEREIRA, J. C. **Métodos de amostragem e tamanho de amostra para avaliar o estado de maturação da uva niágara rosada.** Revista Brasileira de Biometria, São Paulo, v.27, n.4, p.501-518, 2009.

ARIZONO, H.; MATSUOKA, S.; GHELLER, A. C. A.; MATSUDA, Y.; HOFFMANN, H. P.; BASSINELLO, A. I.; MANESES, L. L. **Alternativas para avaliação de produção de cana-de-açúcar.** Stab, v.16, n.5, p.20, 1998.

BALESTRASSI, P. P. e PAIVA, A. P. **Estatística aplicada – notas compiladas.** UNIFEI - Universidade Federal de Itajubá - Instituto de Engenharia de Produção e Gestão. 2007.

BARBIN, D. e MONTALVAN, R. **Estatística e melhoramento genético de plantas.** In: Destro, D. e R. Montalván. (Org.). **Melhoramento genético de plantas.** Editora da Universidade Estadual de Londrina. Londrina, PR, p.85-102, 1999.

BARBOSA, M. H. P.; SILVEIRA, L. C. I. e DONIZETTE, R. **Tamanho da amostra para avaliação de famílias de cana-de-açúcar.** Anais do 7º Congresso STAB, p.57-59, 1999.

BARBOSA, M. H. P. e SILVEIRA, L. C. I. **Metodologias de seleção, progressos e mudanças no programa de melhoramento genético da cana-de-açúcar da Universidade Federal de Viçosa.** STAB, janeiro-fevereiro/2000 – Vol. 18 nº 3, p.30-32.

BARROS, I. e TAVARES, M. **Estimativa do tamanho ótimo de parcelas experimentais através de cálculos algébricos.** Bragantia, Campinas, 54(1): 209-215, 1995.

BRESSIANI, J. A. **Seleção sequencial em cana-de-açúcar**. Tese (doutorado). Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, 2001.

CARGNELUTTI FILHO, A.; RIBEIRO, N. D.; STORCK, L.; JOST, E.; POERSCH, N. L. **Tamanho de amostra de caracteres de cultivares de feijão**. Ciência Rural, Santa Maria, v.38, n.3, p.635-642, mai-jun, 2008.

CARGNELUTTI FILHO, A.; EVANGELISTA, D. H. R.; GONÇALVES, E. C. P.; STORCK, L. **Tamanho de amostra de caracteres de genótipos de soja**. Ciência Rural, Santa Maria, v.39, n.4, p.983-991, jul, 2009.

CARGNELUTTI FILHO, A.; TOEBE, M.; BURIN, C.; SILVEIRA, T. R.; CASAROTTO, G. **Tamanho de amostra para estimação do coeficiente de correlação linear de Pearson entre caracteres de milho**. Pesquisa agropecuária brasileira, Brasília, v.45, n.12, p.1363-1371, dez. 2010.

CARGNELUTTI FILHO, A.; TOEBE, M.; STURZA, V. S.; BOLZAN, A.; SILVEIRA, T. R.; DEQUECH, S. T. B. **Tamanho de amostra para a estimação da média de duração dos períodos larval, pupal e larval mais pupal de *Microtheca* spp. por *bootstrap***. Ciência Rural, Santa Maria, v.42, n.1, p.13-16, jan, 2012.

CARNEIRO, F. F.; ZEVIANI, W. M.; SANTOS, J. B.; CARVALHO, R. S. B.; ALVES, F. C.; DIAS, J. A. **Minimum number of common bean plants per plot to assess field resistance to white mold**. Crop Breeding and Applied Biotechnology 11: 358-364, 2012.

CATAPATTI, T. R.; GONÇALVES, M. C.; SILVA NETO, M. R.; SOBROZA, R. **Tamanho de amostra e número de repetições para avaliação de caracteres agrônômicos em milho pipoca**. Ciência e Agrotecnologia, Lavras, v. 32, n. 3, p. 855-862, maio/jun., 2008.

CESNIK, R. e MIOCQUE, J. **Melhoramento da cana-de-açúcar**. Brasília: Embrapa Informações Tecnológicas, 2004, 307p.

CHEAVEGATTI-GIANOTTO, A. ABREU, H. M. C.; ARRUDA, P.; BESPALHOK FILHO, J. C.; BURNQUIST, W. L.; CRESTE, S.; CIERO, L.; FERRO, J. A.; FIGUEIRA, A. V. O.; FILGUEIRAS, T. S.; GROSSI-DE-SÁ, M. F.; GUZZO, E. C.; HOFFMANN, H. P.; LANDELL, M. G. A.; MACEDO, N.; MATSUOKA, S.; REINACH, F. C.; ROMANO, E.; SILVA, W. J.; SILVA FILHO, M. C.; ULIAN, E. C. **Sugarcane (*Saccharum x officinarum*): A reference study for the regulation of genetically modified cultivars in Brazil.** Tropical Plant Biology.

COCHRAN, W. G. **Sampling techniques.** 3d. ed. New York: Wiley. 1977.

CONAB - Companhia Nacional de Abastecimento. **Acompanhamento da safra brasileira: Cana-de-açúcar, terceiro levantamento, dezembro/2012 - Companhia Nacional de Abastecimento.** Brasília: Conab 2012.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético.** Vol 1. 3 ed. Viçosa: Editora UFV, 2004. 480 p.

EFRON, B. **Bootstrap methods: Another look at the jackknife.** Annals of Statistics, 7: 1-26. 1979.

FEIJÓ, S.; STORK, L.; LÚCIO, A. D.; LOPES, S. J. **Heterogeneidade do solo e de tamanho de amostra antes e após cultivos com abobrinha italiana em estufa plástica.** Ciência Rural, Santa Maria, v.36, n.6, p.1744-1748, nov-dez, 2006.

FERREIRA, D. F. **Recursos computacionais utilizando R.** Universidade Federal de Lavras, Departamento de Ciências Exatas. 2011

FONSECA, J. S. e MARTINS, G. A. **Curso de estatística.** Quinta Edição. Editora ATLAS. São Paulo, SP, 1995.

FOX, J. e WEISBERG, S. **Bootstrapping regression models in R - an appendix to an R companion to applied regression.** In: FOX, J. e WEISBERG, S. **An R companion to applied regression.** Second Edition. SAGE Publications, Inc. 2011.

GHELLER, A. C. A.; MENEZES, L. L.; MATSUOKA, S.; MASUDA, Y.; HOFFMANN, H. P.; ARIZONO, H.; GARCIA, A. A. F. **Manual de método alternativo para medição da produção de cana-de-açúcar**. Araras: UFSCar – CCA – DBV, 1999.

HAESBAERT, F. M.; SANTOS, D.; LÚCIO, A. D.; BENZ, V.; ANTONELLO, B. I.; RIBEIRO, A. L. P. **Tamanho de amostra para experimentos com feijão-de-vagem em diferentes ambientes**. Ciência Rural, Santa Maria, v.41, n.1, p.38-44, jan, 2011.

JAMES, G. **Sugarcane**. Second Edition. Blackwell Science. Oxford, UK. 2004.

KERBAUY, G. B. **Fisiologia vegetal**. Segunda edição expandida, revisada e atualizada. 2a. ed. Rio de Janeiro: Editora Guanabara Koogan Ltda., 2008.

LANDELL, M. G. A.; VASCONCELOS, A. C. M.; SILVA, M. A.; PERECIN, D.; CARVALHO, R. S. R.; BARBOSA, V.; PENNA, M. J. **Validação de métodos de amostragem para a estimativa da produção de cana-de-açúcar, em áreas de colheita mecanizada**. STAB – novembro/dezembro/1999 – Vol. 18 nº 2, p.48-51.

LEITE, M. S. O. **Tamanho da amostra para seleção de famílias de cana-de-açúcar**. Dissertação (mestrado). Universidade Federal de Viçosa, Viçosa, MG, 2007.

LESMAN, K. J. e ATKINS, R. E. **Optimum plot size and relative efficiency of lattice designs for grain sorghum yield tests**. Crop Science, Madison, v.3, p.477-481, 1963.

LÚCIO, A. D.; SOUZA, M. F.; HELDWEIN, A. B.; LIEBERKNECHT, D.; CARPES, R. H.; CARVALHO, M. P. **Tamanho da amostra e método de amostragem para avaliação de características do pimentão em estufa plástica**. Horticultura Brasileira, Brasília, v. 21, n. 2, p. 180-184, abril/junho, 2003.

LOGAN, M. **Biostatistical design and analysis using R – A practical guide**. First Edition. Wiley-Blackwell. Oxford, UK. 2010.

MARODIM, V. S.; STORCK, L.; LOPES, S. J.; SANTOS, O. S.; SCHIMIDT, D. **Delineamento experimental e tamanho de amostra para alface cultivada em hidroponia**. Ciência Rural, Santa Maria, v.30, n.5, p.779-781, 2000.

MATSUOKA, S.; GARCIA, A. A. A.; ARIZONO, H. **Melhoramento da cana-de-açúcar**. In: BORÉM, A. **Melhoramento de espécies cultivadas**. Viçosa: Editora UFV, 1999. p. 205-251.

MEIER, V.D. e LESSMAN, K.J. **Estimation of optimum field plot shape and size testing yield in *Crambe abyssinica* HORDNT**. Crop Science, v.11, p.648-650, 1971.

MOURA, K. H. S. **Determinação do tamanho da amostra para avaliação de híbridos de melão amarelo**. Dissertação (mestrado). Universidade Federal Rural do Semi-árido, Mossoró, RN, 2018.

NUNES, G. H. S.; TORQUATO, J. E.; SALES JÚNIOR, R.; FERREIRA, H. A.; BEZERRA NETO, F. **Tamanho amostral para estimar o teor de sólidos solúveis totais em talhões de melão amarelo**. Caatinga (Mossoró,Brasil), v.19, n.2, p.117-122, abril/junho 2006.

OLIVEIRA, G. M. V. **Tamanho de parcelas experimentais para *Eremanthus erythropappus***. Dissertação (mestrado). Universidade Federal de Lavras, Lavras, MG, 2011.

OLIVEIRA, G. M. V.; MELLO, J. M.; LIMA, R. R.; SCOLFORO, J. R. S.; OLIVEIRA, A. D. **Tamanho e forma de parcelas experimentais para *Eremanthus erythropappus***. Cerne, Lavras, v. 17, n. 3, p. 327-338, jul./set. 2011.

OLIVEIRA, R. A. **Seleção de famílias de maturação precoce em cana-de-açúcar via Reml/Blup**. Tese (doutorado). Universidade Federal do Paraná, Curitiba, PR, 2007.

OYEYEMI, G. M. **Comparison of bootstrap and jackknife methods of resampling in estimating population parameters**. Global Journal of Pure and Applied Sciences, Vol 14, No. 2, 2008: 217-220.

PARANAÍBA, P. F.; FERREIRA, D. F.; MORAIS, A. R. **Tamanho ótimo de parcelas experimentais: Proposição de métodos de estimação.** Revista Brasileira de Biometria, São Paulo, v.27, n.2, p.255-268, 2009.

PEQUENO, S. A.; PINHEIRO, J. B.; ZUCCHI, M. I.; VENCONVSKY, R.; COELHO, A. S. G.; TRINDADE, M. G. **determinação do número de marcadores RAPD para estudos da diversidade genética em soja utilizando o método *bootstrap*.** Bioscience Journal, Uberlândia, v. 19, n. 2, p. 45-48, May/Aug. 2003.

PÉLLICO NETO, S. e BRENA, D. A. **Inventário florestal.** Curitiba: UFPR. 1997.

SILVA, P. S. L.; SILVA, G. S.; RODRIGUES, M. V.; MACHADO, A. A. **Métodos de amostragem e tamanho da amostra para alguns caracteres do milho.** Ciência Agronômica, Fortaleza, 24 (1/2), pág 5-10, Junho/Dezembro. 1993.

SILVA, D. N. **O Método bootstrap e aplicações à regressão múltipla.** Dissertação (mestrado). Universidade Estadual de Campinas, Campinas, SP, 1995.

SILVA, J. A. A.; CAMPOS, A. L.; VEIGA, A. F. S. L.; MOREIRA, A. F. C.; MARQUES, E. J. **Estimativa da suficiência amostral para avaliar intensidade de infestação da *Diatraea* spp. em cana-de-açúcar.** Pesquisa agropecuária brasileira, Brasília, v.32, n. 10, p. 1003-1007, out. 1997.

SILVA, J.; LIMA E SILVA, P. S.; NUNES, R. P. **Amostragem e tamanho da amostra na estimação de caracteres da espiga do milho.** Pesquisa agropecuária brasileira, Brasília, v.33, n.12, p.1983-1988, dez. 1998.

SILVA, J. G. C. **Estatística experimental: Planejamento de experimentos (versão preliminar).** Universidade Federal de Pelotas, Instituto de Física e Matemática, Departamento de Matemática e Estatística. Pelotas, RS. 2007.

SMITH, H. F. **An empirical law describing heterogeneity in the yields of agricultural crops.** Journal of Agricultural Science 28: 1-23. 1938.

STORCK, L. GARCIA, D. C., LOPES, S. J., ESTEFANEL, V. **Experimentação vegetal**. Primeira Edição. Editora UFSM (Universidade Federal de Santa Maria). Santa Maria, RS, 2000.

STORCK, L. GARCIA, D. C., LOPES, S. J., ESTEFANEL, V. **Experimentação Vegetal**. Segunda Edição. Editora UFSM (Universidade Federal de Santa Maria). Santa Maria, RS, 2006.

TOEBE, M.; BOTH, V.; CARGNELUTTI FILHO, A; BRACKMANN, A.; STORCK, L. **Dimensionamento amostral para avaliar firmeza de polpa e cor da epiderme em pêssgo e maçã**. Ciência Agronômica, v. 42, n. 4, p. 1026-1035, out-dez, 2011.

XIE, C. e MOSJIDIS, J. A. **Influence of sample size on precision of heritability and expected selection response in red clover**. Plant Breeding 116, 83-88 (1997).

ZEVIANI, W. M. **Avaliação de modelos de regressão não linear na cinética de liberação de potássio de resíduos orgânicos**. Dissertação (mestrado). Universidade Federal de Lavras, Lavras, MG, 2009.

ANEXO 1 – QUADROS DE ANOVA RESUMIDOS

ANOVA do experimento a campo (considerando-se apenas os quatro clones efetivamente colhidos).

FV	GL	F	p-valor	Clone	Médias (TCH)
Blocos	2	0,104	0,902 ns	RB867515	122,34
Tratamentos	3	1,129	0,409 ns	RB005151	139,01
		ns (não significativo)		RB006655	124,75
				RB006638	113,82

Nota: ANOVA efetuada com os dados da colheita total das parcelas experimentais.

ANOVA-R: ANOVA real com clone testador.

FV	GL	F	p-valor	Clone	Médias (TCH)
Blocos	2	0,273	0,768 ns	RB867515	122,34 a
Tratamentos	4	6,832	0,011 **	RB005151	139,01 a
		ns (não significativo)		RB006655	124,75 a
		** (significativo para $\alpha = 0,05$)		RB006638	113,82 ab
				Testador	77,46 b

Nota: ANOVA efetuada com os dados da colheita total das parcelas experimentais e das três parcelas do clone testador, cujos dados foram simulados no ambiente estatístico “R”. Médias seguidas de letras iguais não diferem umas das outras pelo teste de Tukey a 5% de probabilidade de erro.

ANEXO 2 – ALGUNS ALGORITMOS DE PROGRAMAÇÃO UTILIZADOS NO AMBIENTE COMPUTACIONAL “R”

```
#=====
# Bootstrap em colmos de cana - Para a estatística "CV" da variável MC
# Parcela: 4 linhas de 8 metros
#=====
#-----
# definições da sessão
#-----
require(lattice)
require(latticeExtra)
require(plyr)
#-----
# leitura de dados
#-----
da <- read.table("ex1.txt", header=TRUE, sep="\t", dec=".")
str(da)
levels(da$GEN)[5] <- "RB.867515"
da$parcela <- with(da, interaction(GEN, BLOCO, drop=TRUE))
#-----
# avaliações da distribuição e da normalidade dos dados
#-----
qqmath(~PC|parcela, data=da)
histogram(~PC|parcela, data=da)
densityplot(~PC|parcela, data=da)
#-----
# conferência do número de dados
#-----
xtabs(~parcela, da)
#-----
# configurações dos bootstraps
#-----
levels(da$parcela)
nlevels(da$parcela)
db <- split(da$PC, f=da$parcela)
str(db)
ngrid <- c(1, 2,3,5,10,15,20,25,30,35,40,45,50,60,70,80,
          90,100,125,150,200,250,507) # tamanhos de amostras usados
length(ngrid)
ncolm <- 1000          # número de amostras bootstrap
res <- lapply(db,
  function(p){
    lapply(ngrid, function(n){
      matrix(replicate(ncolm, sample(p, n, replace=TRUE)), ncol=n)
    })
  })
str(res)
cv <- function(x) 100*sd(x)/mean(x)
cv.cal <- lapply(res,
  function(p){
    lapply(p, function(n) apply(n, MARGIN=1, mean))
  })
str(cv.cal)
cv.cal2 <- lapply(cv.cal,
  function(perc){
    stat <- do.call(c, perc)
    data.frame(ngrid=rep(ngrid, each=ncolm), stat=stat)
  })
str(cv.cal2)
cv.cal3 <- ldply(cv.cal2)
str(cv.cal3)
cv.cal3$parcela <- factor(cv.cal3$id, levels=levels(da$parcela))
cv.cal4 <- with(cv.cal3,
  aggregate(cbind(stat=stat), list(parcela=parcela, ngrid=ngrid), cv))
#-----
# gráficos da estatística
#-----
xyplot(stat~ngrid|parcela, data=cv.cal3, xlab="TAMANHO AMOSTRAL", ylab="CV",
  main="TAMANHO AMOSTRAL x CV", type=c("p","smooth"))
```

```

xyplot(stat~ngrid|parcela, data=cv.cal4, xlab="TAMANHO AMOSTRAL", ylab="CV",
main="TAMANHO AMOSTRAL x CV", type=c("p","smooth"))
#-----
# ajustes - aplicando os modelos a todas as parcelas
#-----
sub <- subset(cv.cal4, parcela==levels(da$parcela)[1] & ngrid)
cv.cal5 <- subset(cv.cal4, ngrid>=1)
#-----
# resposta platô linearizada
#-----
nls.aj <- by(cv.cal5, INDICES=cv.cal5$parcela,
FUN=nls,
formula=stat~A+B*ngrid*(ngrid<=x0)+B*x0*(ngrid>x0),
start=list(A=15, B=-0.2, x0=50))
sapply(nls.aj, coef)
pred <- data.frame(ngrid=seq(1,max(ngrid),by=1))
aux <- lapply(nls.aj,
function(i){
cbind(pred, y=predict(i, newdata=pred))
})
aux <- ldply(aux)
str(aux)
aux$parcela <- factor(aux$.id, levels=levels(da$parcela))
xyplot(stat~ngrid|parcela, data=cv.cal4, xlab="Tamanho da amostra (nr. de colmos)", ylab="CV (%)", main="Modelo
Linear Segmentado com Resposta Platô - MC")+as.layer(xyplot(y~ngrid|parcela, data=aux, type="l"))
#-----
# quadrático com resposta platô
#-----
nls.aj <- by(cv.cal5, INDICES=cv.cal5$parcela, # modelo 3a
FUN=nls,
formula=stat~(h0+h1*ngrid+h2*(ngrid^2))*(ngrid<=-0.5*h1/h2)
+(h0+1*(-h1^2/(4*h2)))*(ngrid>-0.5*h1/h2),
start=list(h0=100,h1=-9,h2=0.15))
sapply(nls.aj, coef)
pred <- data.frame(ngrid=seq(1,max(ngrid),by=1))
pred <- data.frame(ngrid=seq(1,max(ngrid),by=1))
aux <- lapply(nls.aj,
function(i){
cbind(pred, y=predict(i, newdata=pred))
})
aux <- ldply(aux)
str(aux)
aux$parcela <- factor(aux$.id, levels=levels(da$parcela))
xyplot(stat~ngrid|parcela, data=cv.cal4, xlab="Tamanho da amostra
(nr. de colmos)", ylab="CV (%)", main="Modelo Quadrático Segmentado com Resposta Platô -
MC")+as.layer(xyplot(y~ngrid|parcela, data=aux, type="l"))
xabs<-function(x){-0.5*coef(x)[2]/coef(x)[3]}
XABS<-ldply(nls.aj,xabs);XABS
write.table(XABS,"modelo_qrp.xls",sep="\t",dec=",")
#-----
# modelo do CV relativo
#-----
q <- 0.15
nls.aj <- by(cv.cal5, INDICES=cv.cal5$parcela,
FUN=nls,
formula=stat~b0+(b1-b0)/(1+((1-q)/q)*(ngrid-1)/(b2-1)),
start=list(b0=1, b1=50, b2=20))
sapply(nls.aj, coef)
pred <- data.frame(ngrid=seq(1,max(ngrid),by=1))
aux <- lapply(nls.aj,
function(i){
cbind(pred, y=predict(i, newdata=pred))
})
aux <- ldply(aux)
str(aux)
aux$parcela <- factor(aux$.id, levels=levels(da$parcela))
xyplot(stat~ngrid|parcela, data=cv.cal4, xlab="Tamanho da amostra (nr. de colmos)", ylab="CV (%)", main="Modelo do
CV Relativo - MC")+as.layer(xyplot(y~(ngrid+1)|parcela, data=aux, type="l"))
#-----
# curva potencial
#-----
nls.aj <- by(cv.cal5, INDICES=cv.cal5$parcela,
FUN=nls,
formula=stat~Z/ngrid^W,
start=list(Z=35,W=-0.5))
sapply(nls.aj, coef)

```

```

pred <- data.frame(ngrid=seq(1,max(ngrid),by=1))
aux <- lapply(nls.aj,
  function(i){
    cbind(pred, y=predict(i, newdata=pred))
  })
aux <- ldply(aux)
str(aux)
aux$parcela <- factor(aux$id, levels=levels(da$parcela))
xyplot(stat~ngrid|parcela, data=cv.cal4, xlab="Tamanho da amostra
(nr. de colmos)", ylab="CV (%)", main="Modelo da Curva Potencial - MC")+as.layer(xyplot(y~ngrid|parcela, data=aux,
type="l"))

#=====
# Mil simulações bootstrap para ANOVA (TCH)
#=====
#-----
# Definições da sessão.
#-----
require(lattice)
#-----
# Dados.
#-----
db <- read.table("ex2.txt", header=TRUE, sep="\t", dec=".")
names(db) <- tolower(names(db))
str(db)
db <- transform(db, bloco=factor(bloco),
  parcela=interaction(gen, bloco, drop=TRUE))
str(db)
#-----
# Análise de variância.
#-----
m0 <- aov(tch~bloco+gen+Error(parcela), data=db)
summary(m0)
#-----
# Procedimento:
#-----
## A reamostragem bootstrap pegará "n" colmos dentro de cada parcela.
## Para cada sorteio, realiza-se uma análise de variância e armazena-se
## o valor do CV experimental, de F e p-valor, além do ranking dos clones.
## Este procedimento é repetido "B" vezes para cada "n". Ao final,
## ajustam-se modelos que expliquem o comportamento das estatísticas
## armazenadas, em função de "n".
#-----
# Ordenar os dados para facilitar a implementação do procedimento.
#-----
db <- db[with(db, order(gen, bloco)),]
head(db)
#-----
# Lista com os índices dos registros dentro de cada parcela.
#-----
db$i <- seq_along(db$tch)
L <- split(db$i, db$parcela)
sapply(L, length) # número de colmos em cada parcela.
#-----
# Função que retira uma amostra.
#-----
retiraramostra <- function(L, n, B){
  replicate(n=B, expr=c(sapply(L,
    function(x){
      sample(x, n, replace=TRUE)
    })))
}
x <- retiraramostra(L, n=2, B=4);x # testando a função recém criada
with(db[x[,1,],, tapply(tch, parcela, length)) # número de registros dentro da parcela
#-----
# Função que realiza a ANOVA e armazena os p-valores e as médias.
#-----
pegarvalores <- function(formula=tch~bloco+gen+Error(parcela), data, sel){
  m0 <- aov(formula, data=data[sel,])
  m <- model.tables(m0, type="means")
  mc <- m[[1]][[3]] # médias dos clones
  s0 <- summary(m0)
  c(p=s0[[1]][[1]][2,5],
    mc=mc)
}

```

```

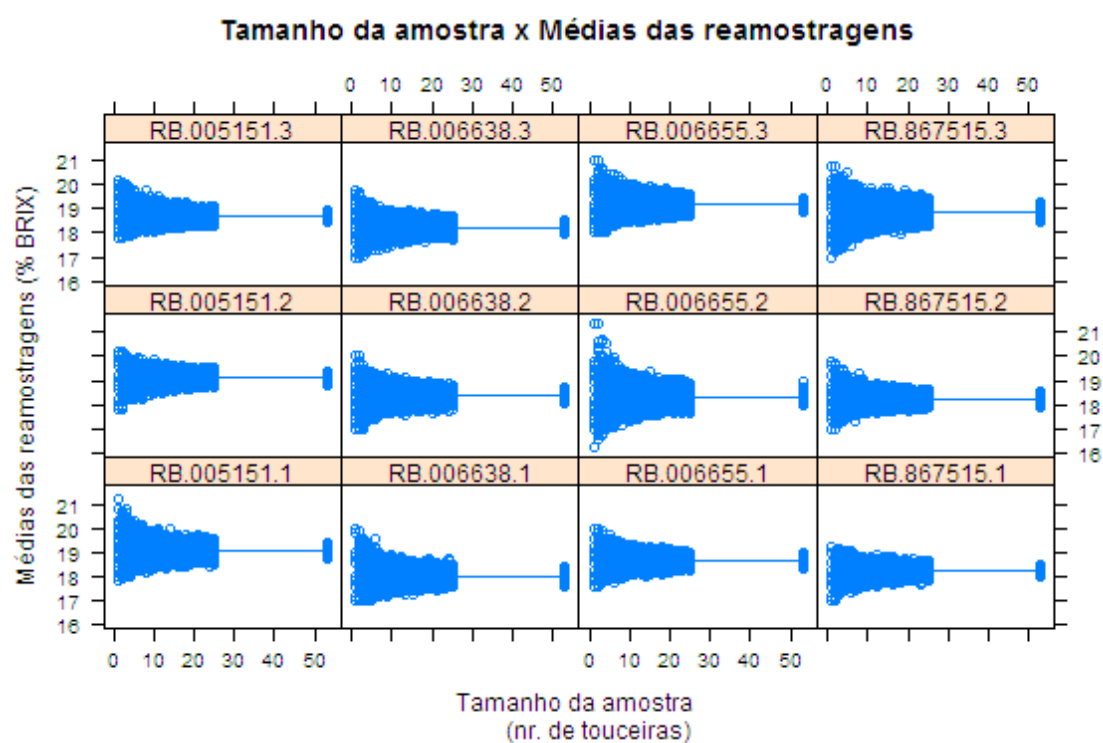
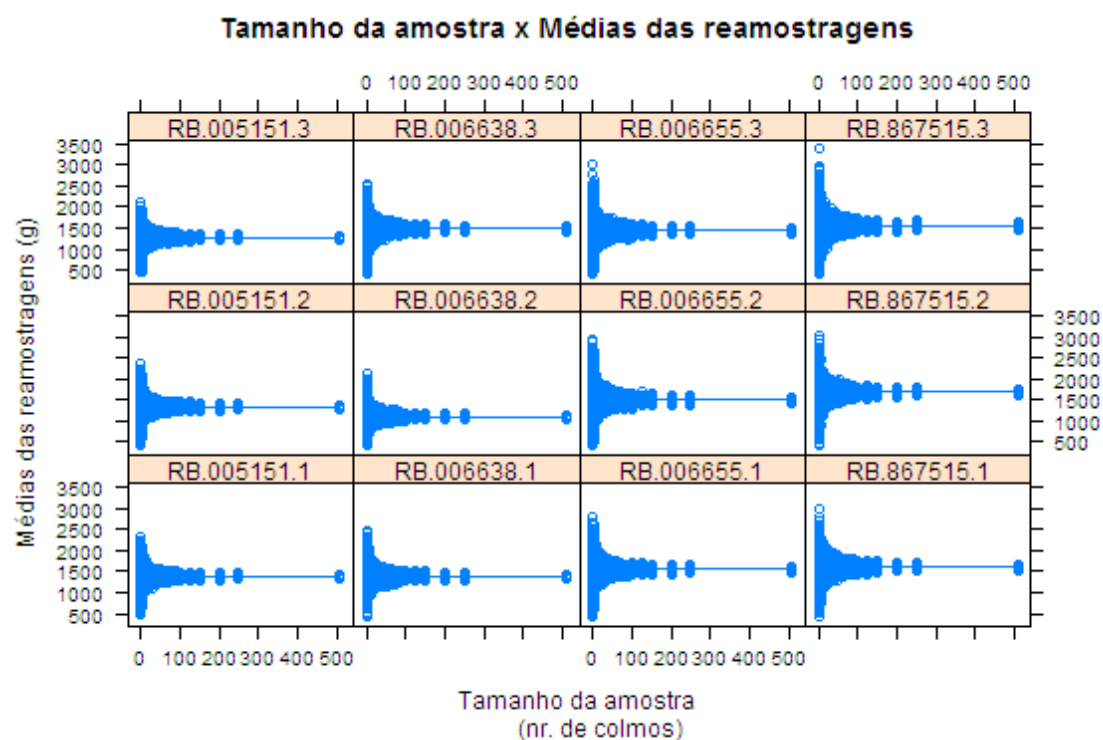
range(x[,1])
db[x[,1],]
pegarvalores(data=db, sel=x[,1])
#-----
# Implementação do procedimento.
#-----
# Sequência de valores de "n" para número de colmos por parcela (ncp)
ncp <- c(1,2,3,5,10,15,20,25,30,35,40,45,50,60,70,80,90,100,125,150,
        200,250,507)
B <- 1000
res <- lapply(ncp,
  function(n){
    x <- retiramostra(L, n=n, B=B)
    y <- apply(x, 2,
      function(i){
        pegarvalores(data=db, sel=i)
      })
    cbind(ncp=n, t(y))
  })
res <- as.data.frame(do.call(rbind, res))
str(res)
#-----
# Gráficos resultantes.
#-----
xyplot(p~ncp, data=res, type=c("p","a"),
  xlab="Tamanho da amostra (nr. de colmos por parcela)",
  ylab="P-valor do teste F",
  main="Simulações: P-valor em função do tamanho de amostra",
  family="serif",font=2)
#-----
# Verificação da proporção de vezes em que o p-valor foi significativo
# para cada tamanho de amostra (ncp).
#-----
str(res)
f <- with(res, aggregate(cbind(p=p), list(ncp=ncp), function(i) sum(i<=0.05)/length(i)))
f
percent<-f$p*100
par(mfrow=c(1,1),family="serif",font=2)
plot(percent~ncp, data=f, type="o",
  xlab="Tamanho da amostra (nr. de colmos por parcela)",
  ylab="P-valores significativos (%)",
  main="Simulações: Análise dos p-valores encontrados",
  font.main=1,cex=1.4,cex.axis=1.4,cex.lab=1.1,cex.main=1.4,cex.sub=1.1)
#-----
# Obter o ranqueamento dos clones.
#-----
RB005151<-c(res$mc.RB.005151) # clone B
RB006638<-c(res$mc.RB.006638) # clone D
RB006655<-c(res$mc.RB.006655) # clone C
RB867515<-c(res$mc.RB.867515) # clone A
CLTESTER<-c(res$mc.TESTER) # clone E
tabela<-data.frame(RB005151,RB006638,RB006655,RB867515,CLTESTER)
#-----
# Função para ordenar os clones.
ordenar<-function(m){
  m[with(m, order(-newX...i.,todos)),]
}
#-----
# Função para extrair os rankings.
#-----
tirarlinha<-function(df){
  nr<-seq(1:(B*length(ncp)))
  md1<-apply(subset(df[nr,]),1,c)
  md2<-apply(subset(md1[1:5,nr]),2,data.frame)
  todos<-c("B","D","C","A","E")
  md3<-lapply(md2,cbind,todos)
  #md3<-lapply(md2,listar)
  #md4<-ldply(md3,c)
  md4<-lapply(md3,ordenar)
  #rk<-apply(md,2,sort)
  #rk
}
rankings1<-tirarlinha(tabela)
rankings2<-lapply(rankings2,c) # objeto com todos os rankings, para todos os
# tamanhos de amostra testados
# agora, bastará separá-los e logo avaliá-los

```

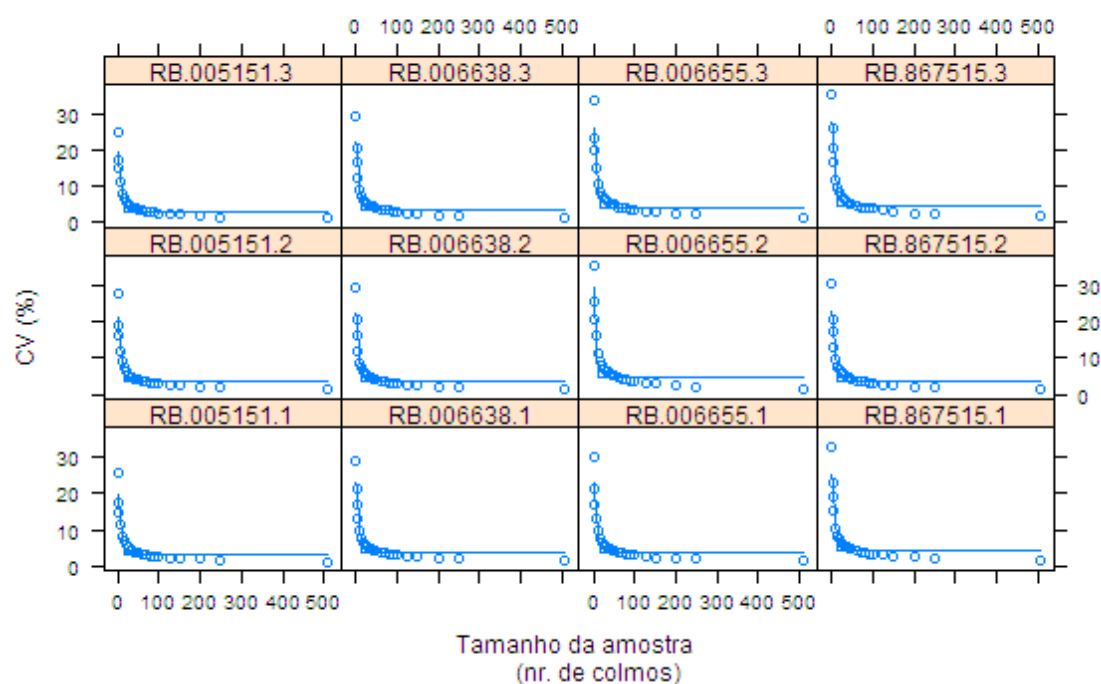
sugestão: imprimir separadamente em pdf, para que se use, então, a ferramenta "search" do leitor, a fim de avaliá-los

#-----

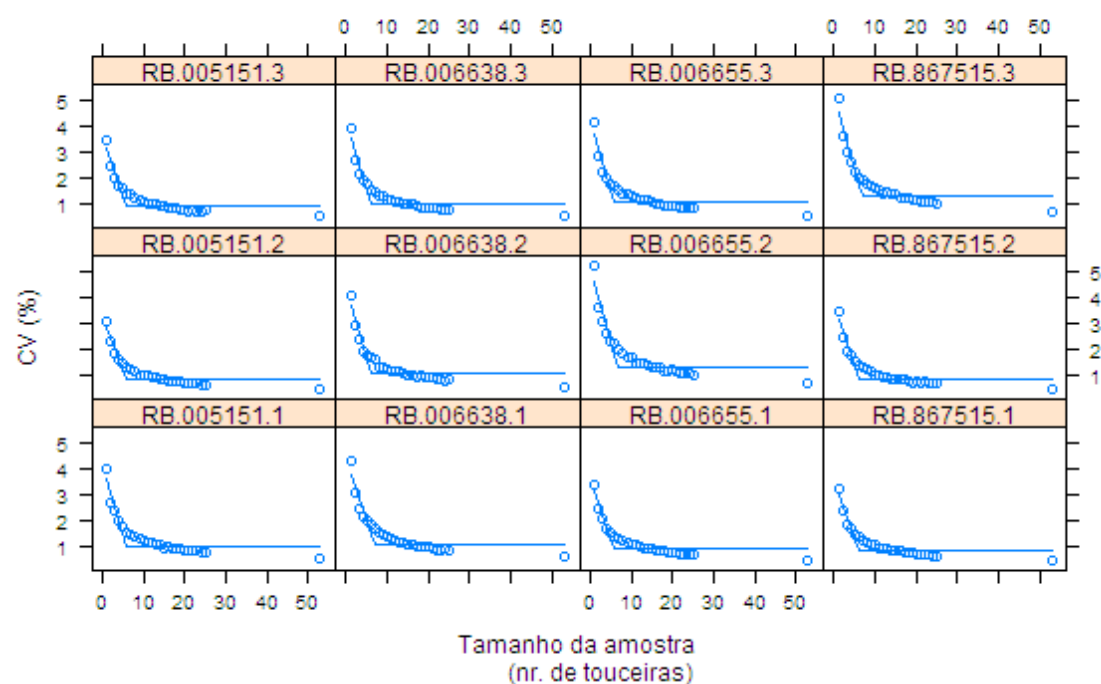
ANEXO 3 – GRÁFICOS RESULTANTES DOS *BOOTSTRAPS* E DOS AJUSTES DOS MODELOS A TODAS AS PARCELAS



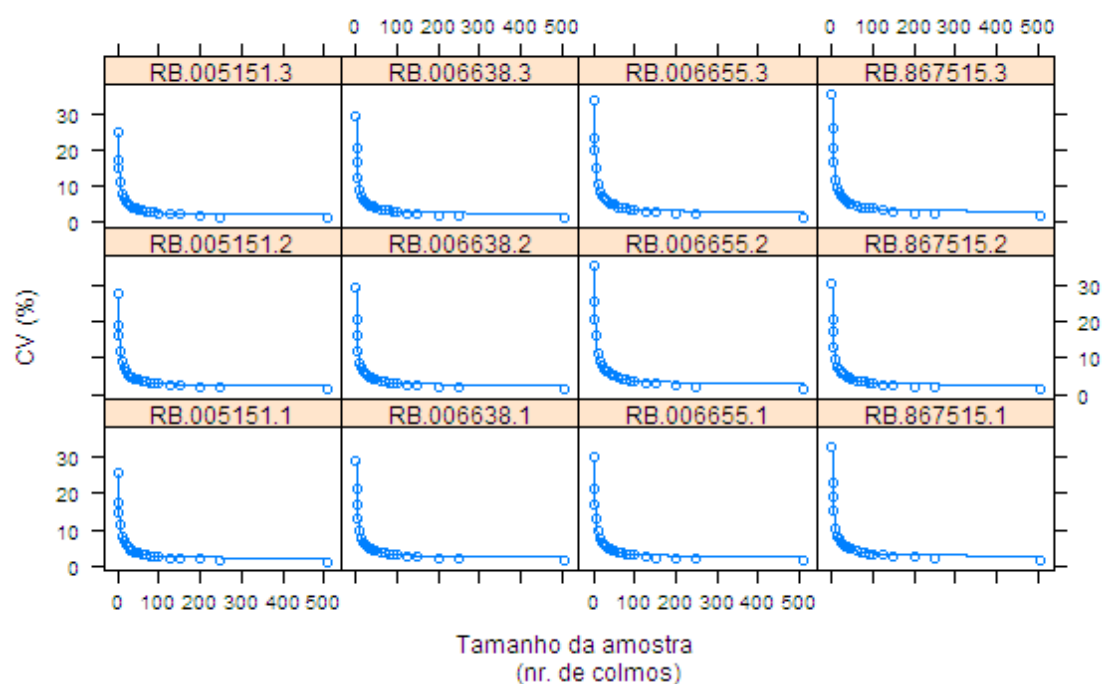
Modelo Linear Segmentado com Resposta Platô - MC



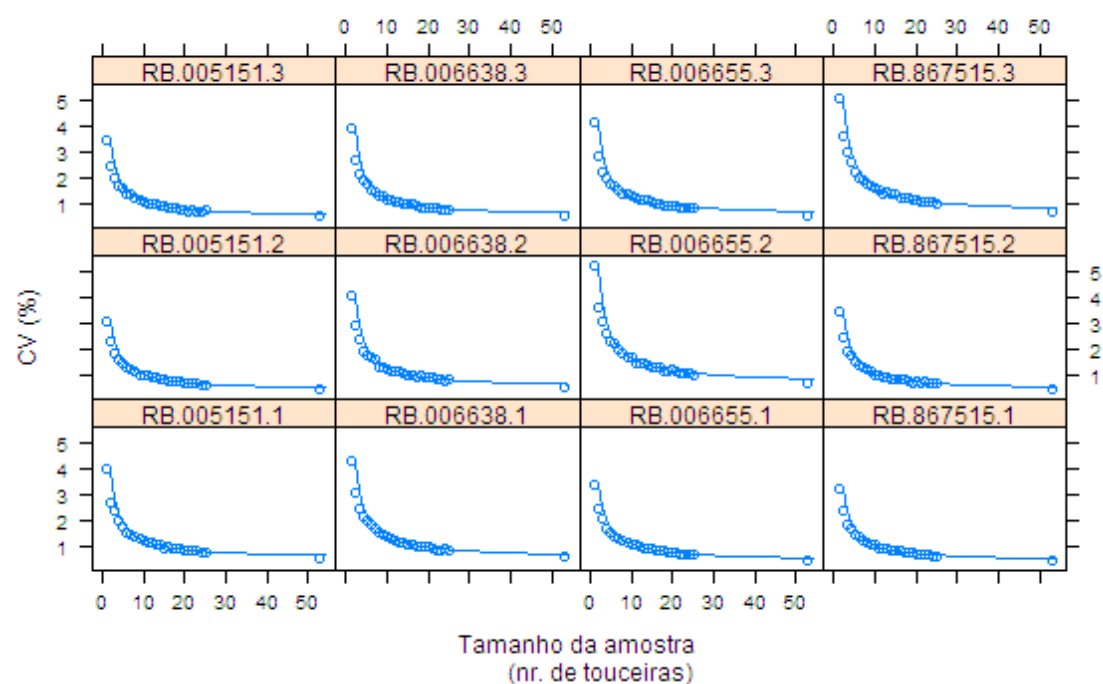
Modelo Linear Segmentado com Resposta Platô - BRIX



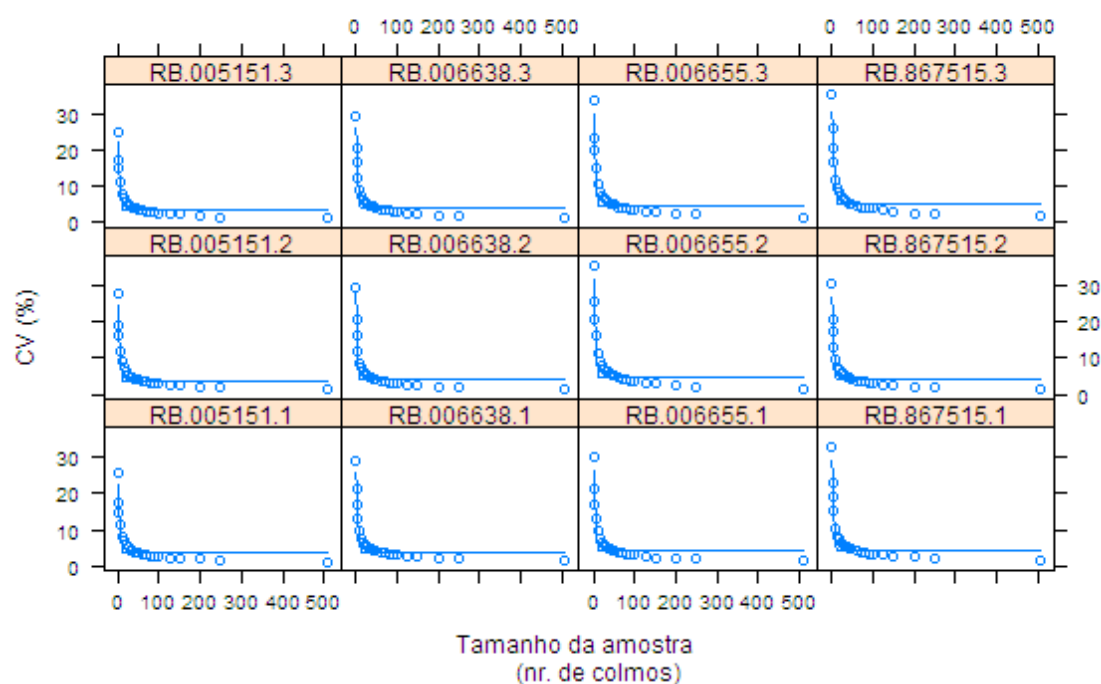
Modelo do CV Relativo - MC



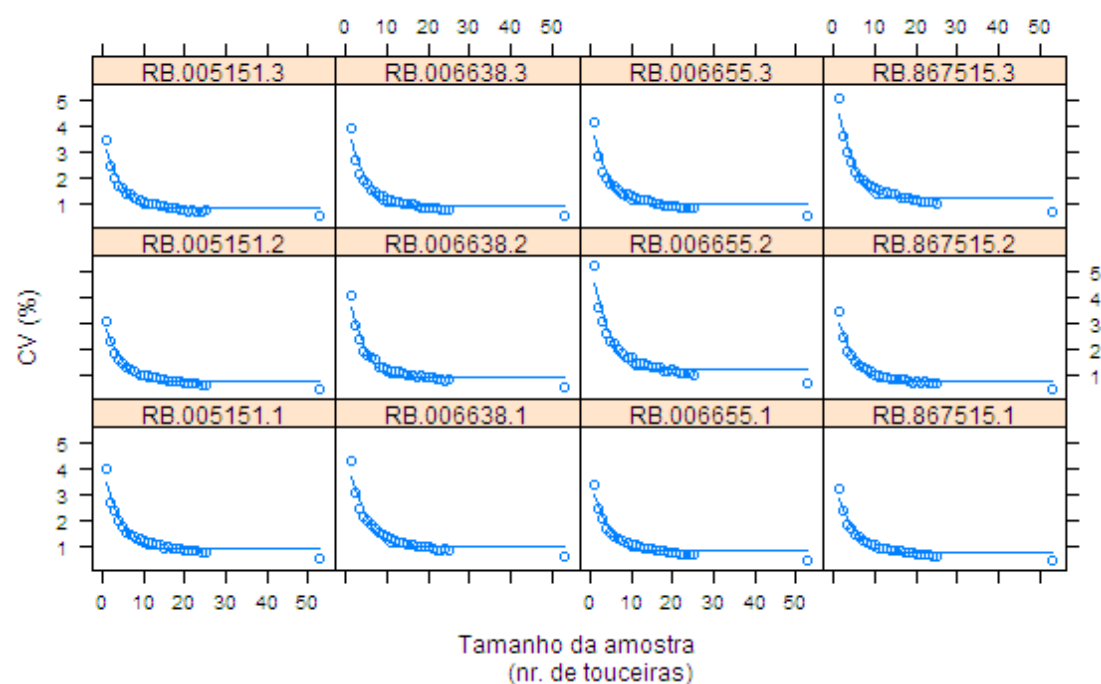
Modelo do CV Relativo - BRIX



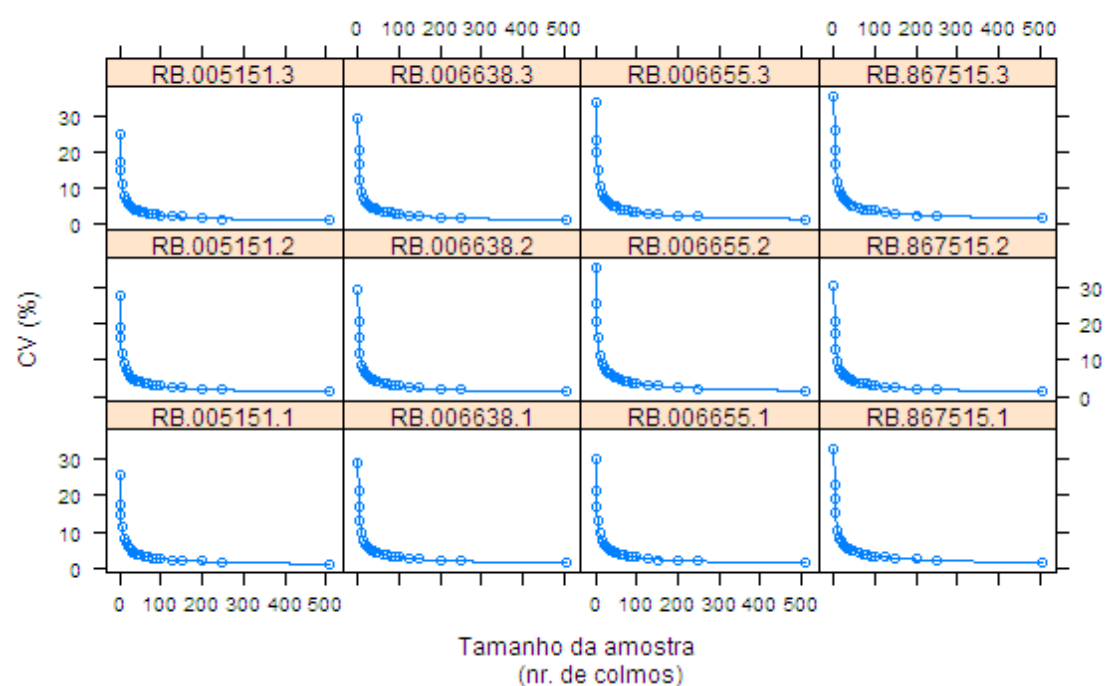
Modelo Quadrático Segmentado com Resposta Platô - MC



Modelo Quadrático Segmentado com Resposta Platô - BRIX



Modelo da Curva Potencial - MC



Modelo da Curva Potencial - BRIX

